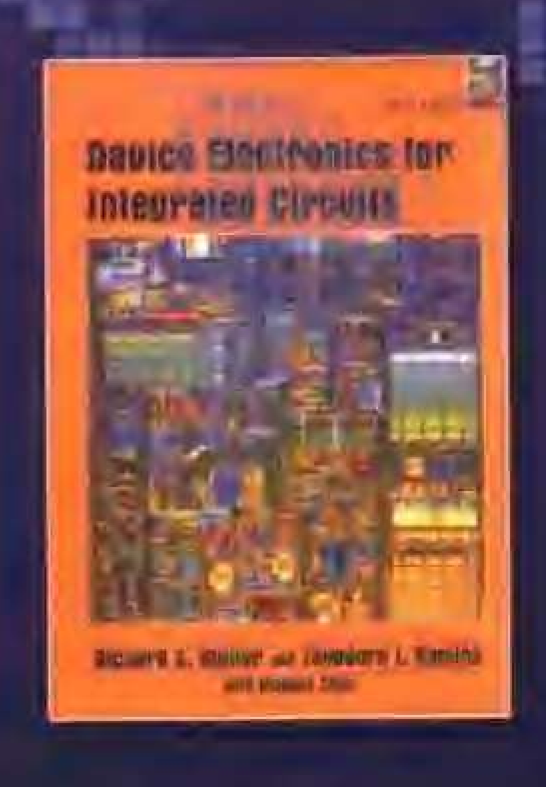
国外电子与通信教材系列

集成电路器件电子学

（第三版）

Device Electronics for Integrated Circuit

Third Edition



Richard S. Muller
[美] Theodore I. Kamins 著
Mansun Chan
王 燕 张 莉 译
许 军 审校





電子工業出版社
Publishing House of Electronics Industry
http://www.phei.com.cn

集成电路器件电子学（第三版）

Device Electronics for Integrated Circuit, Third Edition

当今复杂的微处理器中数百万个电子器件如何协同工作？

控制这些器件特性的物理原理是什么？

集成电路的发明到现在已经有四十多年，目前进入了一个令人振奋但又具有挑战性的阶段。器件的不断缩小似乎已经逼近与结构和材料本身有关的基本极限。

为了了解集成电路的发展，并为这个领域作出贡献，我们需要了解基本的物理原理，并探索其应用领域。

本书是 Richard Muller，Ted Kamins 和 Mansun Chan 编写的《集成电路器件电子学》的第三版。本书力图将例题与实际中丰富多彩的集成电路联系起来。这一独特的方式，即对电学特性的讨论与实际应用的介绍交替进行，有助于读者从整体上了解集成电路，同时也与先进器件模型的发展联系了起来。

本书特点

- 包括 Gauss 定理最新的和扩展的讨论，沟道非常短的 MOS 晶体管，异质结构和异质结晶体管以及工艺模拟
- IC 工艺的最新发展趋势
- 在例题中将基本理论与器件应用联系起来，体现了从理论到应用的途径
- 集中于硅器件进行深入的讨论
- 对例题进行了详细的讨论和说明，有助于读者理解解决实际问题的有效方法
- 提供了丰富的不同难度和复杂程度的习题材料，为学生的练习和导师的作业布置提供了多种选择
- 基于作者在工业界和实验室的实际工作经验，包括了一些相关的主题和例题









责任编辑：杨丽娟
封面设计：毛惠庚

ISBN 7-121-00526-3　　定价：46.00 元

国外电子与通信教材系列

集成电路器件电子学

（第三版）

Device Electronics for Integrated Circuit

Third Edition

Richard S. Muller
【美】 Theodore I. Kamins 著
Mansun Chan
王 燕 张 莉 译
许 军 审校

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内容简介

本书系统介绍集成电路器件电子学，其内容可分为两部分：第一部分是学习半导体器件必需的知识，包括半导体物理和工艺的基本知识，以及金属-半导体接触和pn结理论；第二部分系统深入地阐述了双极晶体管和MOS场效应晶体管的工作原理和特性。书中每章有大量的习题，供读者加深理解所学的知识。

本书可作为高等学校微电子专业本科生和研究生相应课程的教科书或参考书，也可供在相关领域工作的专业技术人员参考。

版权贸易合同登记号 图字：01-2003-3677

图书在版编目(CIP)数据

集成电路器件电子学：第三版/(美)米勒(Muller, R. S.)等著；王燕等译. —北京：电子工业出版社，2004.11

(国外电子与通信教材系列)

书名原文：Device Electronics for Integrated Circuits, Third Edition

ISBN 7-121-00526-3

Ⅰ.集… Ⅱ.①米…②王… Ⅲ.集成电路—电子器件—教材 Ⅳ.TN4

中国版本图书馆CIP数据核字(2004)第110460号

责任编辑：杨丽娟　特约编辑：王银彪

印　　刷：北京天宇星印刷厂

出版发行：电子工业出版社

北京市海淀区万寿路173信箱　邮编100036

经　　销：各地新华书店

开　　本：787×1092　1/16　印张：31.5　字数：806千字

印　　次：2004年11月第1次印刷

印　　数：4 000册　　定价：46.00元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系。联系电话：(010)68279077。质量投诉请发邮件至zlts@phei.com.cn，盗版侵权举报请发邮件至dbqq@phei.com.cn。

序

2001年7月间，电子工业出版社的领导同志邀请各高校十几位通信领域方面的老师，商量引进国外教材问题。与会同志对出版社提出的计划十分赞同，大家认为，这对我国通信事业、特别是对高等院校通信学科的教学工作会很有好处。

教材建设是高校教学建设的主要内容之一。编写、出版一本好的教材，意味着开设了一门好的课程，甚至可能预示着一个崭新学科的诞生。20世纪40年代MIT林肯实验室出版的一套28本雷达丛书，对近代电子学科、特别是对雷达技术的推动作用，就是一个很好的例子。

我国领导部门对教材建设一直非常重视。20世纪80年代，在原教委教材编审委员会的领导下，汇集了高等院校几百位富有教学经验的专家，编写、出版了一大批教材；很多院校还根据学校的特点和需要，陆续编写了大量的讲义和参考书。这些教材对高校的教学工作发挥了极好的作用。近年来，随着教学改革不断深入和科学技术的飞速进步，有的教材内容已比较陈旧、落后，难以适应教学的要求，特别是在电子学和通信技术发展神速、可以讲是日新月异的今天，如何适应这种情况，更是一个必须认真考虑的问题。解决这个问题，除了依靠高校的老师和专家撰写新的符合要求的教科书外，引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，是会有好处的。

一年多来，电子工业出版社为此做了很多工作。他们成立了一个"国外电子与通信教材系列"项目组，选派了富有经验的业务骨干负责有关工作，收集了230余种通信教材和参考书的详细资料，调来了100余种原版教材样书，依靠由20余位专家组成的出版委员会，从中精选了40多种，内容丰富，覆盖了电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等方面，既可作为通信专业本科生和研究生的教学用书，也可作为有关专业人员的参考材料。此外，这批教材，有的翻译为中文，还有部分教材直接影印出版，以供教师用英语直接授课。希望这些教材的引进和出版对高校通信教学和教材改革能起一定作用。

在这里，我还要感谢参加工作的各位教授、专家、老师与参加翻译、编辑和出版的同志们。各位专家认真负责、严谨细致、不辞辛劳、不怕琐碎和精益求精的态度，充分体现了中国教育工作者和出版工作者的良好美德。

随着我国经济建设的发展和科学技术的不断进步，对高校教学工作会不断提出新的要求和希望。我想，无论如何，要做好引进国外教材的工作，一定要联系我国的实际。教材和学术专著不同，既要注意科学性、学术性，也要重视可读性，要深入浅出，便于读者自学；引进的教材要适应高校教学改革的需要，针对目前一些教材内容较为陈旧的问题，有目的地引进一些先进的和正在发展中的交叉学科的参考书；要与国内出版的教材相配套，安排好出版英文原版教材和翻译教材的比例。我们努力使这套教材能尽量满足上述要求，希望它们能放在学生们的课桌上，发挥一定的作用。

最后，预祝"国外电子与通信教材系列"项目取得成功，为我国电子与通信教学和通信产业的发展培土施肥。也恳切希望读者能对这些书籍的不足之处、特别是翻译中存在的问题，提出意见和建议，以便再版时更正。

吴佑寿

中国工程院院士、清华大学教授

"国外电子与通信教材系列"出版委员会主任

出版说明

进入21世纪以来，我国信息产业在生产和科研方面都大大加快了发展速度，并已成为国民经济发展的支柱产业之一。但是，与世界上其他信息产业发达的国家相比，我国在技术开发、教育培训等方面都还存在着较大的差距。特别是在加入WTO后的今天，我国信息产业面临着国外竞争对手的严峻挑战。

作为我国信息产业的专业科技出版社，我们始终关注着全球电子信息技术的发展方向，始终把引进国外优秀电子与通信信息技术教材和专业书籍放在我们工作的重要位置上。在2000年至2001年间，我社先后从世界著名出版公司引进出版了40余种教材，形成了一套"国外计算机科学教材系列"，在全国高校以及科研部门中受到了欢迎和好评，得到了计算机领域的广大教师与科研工作者的充分肯定。

引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，将有助于我国信息产业培养具有国际竞争能力的技术人才，也将有助于我国国内在电子与通信教学工作中掌握和跟踪国际发展水平。根据国内信息产业的现状、教育部《关于"十五"期间普通高等教育教材建设与改革的意见》的指示精神以及高等院校老师们反映的各种意见，我们决定引进"国外电子与通信教材系列"，并随后开展了大量准备工作。此次引进的国外电子与通信教材均来自国际著名出版商，其中影印教材约占一半。教材内容涉及的学科方向包括电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等，其中既有本科专业课程教材，也有研究生课程教材，以适应不同院系、不同专业、不同层次的师生对教材的需求，广大师生可自由选择和自由组合使用。我们还将与国外出版商一起，陆续推出一些教材的教学支持资料，为授课教师提供帮助。

此外，"国外电子与通信教材系列"的引进和出版工作得到了教育部高等教育司的大力支持和帮助，其中的部分引进教材已通过"教育部高等学校电子信息科学与工程类专业教学指导委员会"的审核，并得到教育部高等教育司的批准，纳入了"教育部高等教育司推荐——国外优秀信息科学与技术系列教学用书"。

为做好该系列教材的翻译工作，我们聘请了清华大学、北京大学、北京邮电大学、东南大学、西安交通大学、天津大学、西安电子科技大学、电子科技大学等著名高校的教授和骨干教师参与教材的翻译和审校工作。许多教授在国内电子与通信专业领域享有较高的声望，具有丰富的教学经验，他们的渊博学识从根本上保证了教材的翻译质量和专业学术方面的严格与准确。我们在此对他们的辛勤工作与贡献表示衷心的感谢。此外，对于编辑的选择，我们达到了专业对口；对于从英文原书中发现的错误，我们通过与作者联络、从网上下载勘误表等方式，逐一进行了修订；同时，我们对审校、排版、印制质量进行了严格把关。

今后，我们将进一步加强同各高校教师的密切关系，努力引进更多的国外优秀教材和教学参考书，为我国电子与通信教材达到世界先进水平而努力。由于我们对国内外电子与通信教育的发展仍存在一些认识上的不足，在选题、翻译、出版等方面的工作中还有许多需要改进的地方，恳请广大师生和读者提出批评及建议。

电子工业出版社

教材出版委员会

译 者 序

微电子技术是现代电子信息技术的基础，它深刻影响其他相关学科和工程技术的发展。作为集成电路核心的半导体器件自1947年问世以来，经过半个多世纪的快速发展，已经进入深亚微米量级。了解和掌握集成电路器件的基本理论和物理模型对于微电子及相关专业人员是非常必要的。Richard S. Muller 教授等人所著的《集成电路器件电子学》(Device Electronics for Integrated Circuits)为我们提供了一本深入阐述集成电路器件的物理机制和工作原理并和实践相结合的优秀教材。

作者 Richard S. Muller 教授一直在美国加州大学伯克利分校从事集成电路领域前瞻性的研究与教学工作，在集成电路领域具有相当的知名度和实践经验。合著者 Theodore I. Kamins 教授是加州 Palo Alto 惠普实验室量子科学研究小组的首席科学家，一直从事外延和多晶硅的研究工作。合著者 Mansun Chan 教授任教于香港科技大学，研究领域涵盖了硅器件从深亚微米工艺开发到器件设计、表征和建模。本书内容从半导体物理的基本知识开始，深入介绍现代半导体器件的基本理论和最新进展，并在每一章的最后列出了许多习题和很有价值的参考资料。因此本书可作为高等学校微电子专业本科生和研究生相应课程的教科书或参考书，也可供在相关领域工作的专业技术人员参考。

本书共10章。第1章回顾半导体物理及半导体电学特性，第2章简要介绍现代平面硅工艺技术，第3章介绍金属-半导体接触的电学特性，第4章和第5章介绍同质 pn 结和异质 pn 结的电学特性，第6章和第7章介绍双极晶体管的基本原理和电学特性，第8章介绍 MOS 系统的基本物理知识，第9章和第10章介绍 MOS 场效应晶体管的基本原理和电学特性。

本书在翻译过程中得到了电子工业出版社的热心指导和大力支持，得到了清华大学微电子学研究所多位老师的帮助与指正，包括杨之廉教授、顾祖毅教授、沈延钊教授、陈天鑫教授、周润德教授以及岳瑞峰、潘立阳和邹建平老师，在此深表谢意。虽经仔细校对，但由于译者水平有限，文中定有不当或欠妥之处，望读者批评指正。

译 者

2004年10月于清华园

作者介绍

Richard S. Muller

自20世纪60年代早期硅集成电路发明以来,Richard S. Muller一直从事这一领域的研究工作。他于1962年在加州理工学院获得博士学位。此后加入加州大学伯克利分校,进行前瞻性的研究与教学工作。通过与校内的电路设计人员,器件专家和工艺人员的合作,以及在硅谷几家公司的咨询活动,他在集成电路领域获得了相当的知名度和实践经验。这样的经历为他的新课程的建立提供了坚实的基础。该课程的名称后来被用作本书的书名,并且他已经给伯克利分校的多届学生教授这门课程。70年代后期,他开始从事现在称为"微机电系统"领域的研究,并在1986年与一位同事一起建立了伯克利传感器与执行器中心。他现为美国国家工程院院士,并且是IEEE的会员。其他的一些荣誉包括:IEEE学会Brunetti奖,Stevens工学院复兴奖,Alexander von Humboldt研究奖,伯克利嘉奖,北大西洋公约组织研究员和富尔布赖特研究奖。Muller教授目前还是Stevens工学院董事会成员。

Theodore I. Kamins

加州Palo Alto惠普实验室量子科学研究小组首席科学家。他领导着先进纳米结构的电子材料和器件的研究。他同时也是斯坦福大学电气工程系的顾问教授。他在加州大学伯克利分校获得学位,然后加入了仙童半导体公司的研究与开发实验室。在去惠普之前,他一直从事外延和多晶硅的工作。他是本书的合著者,除此之外,他还是《Polycrystalline Silicon for Integrated Circuits and Display》的作者。他是IEEE和电化学学会的会员,并且获得1989年电化学学会的电子学分会奖。他同时任教于斯坦福大学和加州大学伯克利分校。他还是IEEE Transaction on Electron Devices的副主编,并且为加州大学、牛津大学、IEEE学会、美国真空学会和SEMI教授短期课程。

Mansun Chan

1995年于加州大学伯克利分校获得博士学位,此后任教于香港科技大学。他的研究兴趣涵盖了硅器件领域很广的范围,从深亚微米工艺开发到器件设计、表征和建模。他是SPICE中BSIM3模型的主要贡献者之一,这一模型已经被大多数的美国公司和Compact Model Council作为一个工业标准模型。在2002年1月,他返回加州大学伯克利分校共同指导BSIM计划以及亚70nm器件的下一代器件模型的开发。Mansun在香港科技大学给高年级本科生和研究生教授了六年多的器件和工艺课程,最近也在加州大学伯克利分校授课。在教学工作方面,他已经获得了一些奖项。此外,他还为半导体行业提供公开演讲,为国际会议提供指南。

前　言

随着时间的推移，集成电路(IC)对工程领域以及更广阔的社会生活产生的巨大影响在不断地增加。作为对这种影响的认可，以及对创造集成电路的天才的致谢，Jack Kilby 因其 40 多年前在得克萨斯仪器公司所做的工作被授予 2000 年度*诺贝尔物理学奖。Kilby 创造性地提出了电路集成化制造的概念，并在锗衬底上得到了验证。除此之外，Jean Hoerni 与加州仙童半导体公司的一个研究小组进行了开创性的工作，建立了硅平面工艺，从而开辟出通往现代集成电路的道路。在 20 世纪 60 年代早期，Robert Noyce 与仙童公司的同事们合作制造了第一个硅平面工艺集成电路(见本书卷首插图中的二进制或门电路)，在此基础上发展起来的价值几千亿美元的集成电路产品正主宰着我们现在的数字化世界。

从 Kilby 的创造性概念开始，Hoerni、Noyce 以及他们的合作者创造了总值 1770 亿美元的集成电路全球性市场(2000 年)，支持带动了 1.15 万亿美元的产品市场(2000 年)。集成电路给现代社会打下了明显的印记。最近，Intel 前任首席执行官(CEO) Andrew Grove 先生被时代杂志(TIME)授予 1997 年"年度人物"。Grove 博士是发展集成电路的先驱者，同时他也是 1967 年出版的《Physics and Technology of Semiconductor Devices》一书的作者。这本书第一次描述了平面工艺集成电路的物理、设计及工艺问题。在新千年中，集成电路的影响将超越传统的电子学，延伸到力学、光学、化学和生物学的集成系统中。在这些不同的领域中，集成电路的技术、设计范例、材料以及工艺将得到扩展和丰富，其中形成了所谓的微机电系统(MEMS)。

现代集成电路中，一块芯片上包含千万个有源器件，但是大多数的芯片和 1961 年 Noyce 制造的集成电路一样，都是由硅制造的。20 世纪 80 年代，金属-氧化物-半导体场效应晶体管(MOSFET)取代了先前占主导地位的双极晶体管，并逐渐成为现代半导体器件的主流。过去 20 年中，集成电路领域的工艺发展主要集中在优化 MOSFET 的性能和尽可能地减小它的尺寸。为了突出 MOSFET 的重要性，器件电子学原书的两位作者 Richard Muller 和 Ted Kamins 邀请了在这一领域具有丰富经验的专家 Mansun Chan 教授，对最后两章的内容做了增补和修改。

本书对前一版第 2 章关于集成电路制造工艺的相关内容做了相当大的修改，以适应工艺的发展水平。集成电路复杂性的不断提高以及计算机能力的不断增强，使得计算机辅助工艺和器件模拟成为可能，本书第 2 章和第 5 章对此做了相关的介绍。对高速晶体管的需求，直接导致了异质结双极晶体管(HBT)，尤其是可与常规硅器件集成的硅基 HBT 的广泛应用。有关异质结、HBT 的概念以及更常见的同质结的发展等内容贯穿于本书的第 4 至 7 章。

我们的目标与前一版的"器件电子学"一样，重点在介绍构成集成电路的固态器件中电子的运动规律。在每一章中，我们提出和器件相关的物理问题，详细讨论并说明其基本的物理原理。本书包含的内容较多，不能在一学期课程中全部讲授，教师可以集中于需要着重强调的内容。在导论类的课程中，第 2、7、10 章中的许多内容可以略去不讲。固体物理背景较强的学生

---

* 他与 Z. Alferov 和 H. Kroemer(由于在相关的固态电子学领域，而不是专门在 IC 领域的杰出工作)共同获得诺贝尔物理奖。

可以略去第1章。例题被融入到许多讨论中，各章节末部分习题(标有星号*)的答案在书尾给出。对于一些较为超前的内容以及一些更具挑战性的习题，我们用“†”符号做标记，表明这部分内容可以略过而不会影响本书其他内容的阅读。下面对本书每一章的内容做一个概要性的描述。

第1章　半导体电学特性　本章回顾了半导体物理及半导体电学特性。其中包括：能带理论，掺杂原理，自由载流子统计分布，漂移和扩散。这里只提供了后续章节讨论器件时需要的内容；强调了物理图像而不是数学公式。作为总结，最后介绍了用于专用集成电路的集成霍尔效应磁传感器。附录介绍了求解静电学问题的 Gauss 定理，这是器件分析的重要手段。

第2章　硅工艺　本章简要介绍了现代平面工艺技术，包括硅的晶体生长，晶圆片预处理，氧化，薄膜淀积，晶圆片的扩散掺杂和离子注入掺杂，光刻和图形转移，电极连接和封装。某些章节相当详细，可以用来作为参考。如略去这些章节，整章内容可以在很短的时间内阅读完，但仍能提供给读者关于硅工艺的一个概貌。本章还简单介绍了计算机辅助工艺模拟的概念以及化合物半导体工艺。器件方面介绍了集成电路中的扩散电阻器。

第3章　金属-半导体接触　本章介绍了不同固体材料之间的热平衡。详细描述了 Schottky 势垒接触的电学特性。供选择阅读的小节给出了电流电压关系、Schottky 欧姆接触特性以及金属-半导体接触的表面效应的严格推导。器件应用介绍了实际的 Schottky 二极管。

第4章　pn 结　本章介绍了与不同掺杂类型的半导体有关的电学特性。详细描述了反偏情况下同质 pn 结和异质 pn 结的电学特性。本章还包括有关 pn 结击穿的内容供选择阅读，可以略过这个内容而不影响阅读的连续性。小结讨论了结型场效应晶体管。

第5章　pn 结电流　本章推导了自由载流子的连续性方程。介绍了产生和复合的概念，并用来描述 pn 结在正向偏置下的电学特性。二极管的少子存储和数学模拟为以后晶体管等效电路的发展提供了基础。供选择阅读的小节包括对 Shockley-Hall-Read 复合理论和空间电荷区产生与复合的完整描述。器件方面集中介绍了二极管在集成电路中的应用，并简单介绍了由化合物半导体制备的发光二极管。第5章之后，对 MOS 器件感兴趣的读者可以先阅读第8至10章，然后再回来阅读关于双极晶体管的第6章和第7章，这样并不影响阅读的连贯性。

第6章　双极晶体管Ⅰ：基本特性　本章对双极晶体管的讨论是在电荷控制模型的框架下进行的。以熟悉的同质掺杂晶体管为例来介绍基本的物理原理。这样的处理问题的方法自然地过渡到偏压范围、电流增益以及 Ebers-Moll 模型等主题。双极晶体管的概念被扩展到异质结双极晶体管。器件讨论集中在用于放大电路和开关电路中的平面 IC 晶体管的特性。

第7章　双极晶体管Ⅱ：局限性与模型化　本章的主题面向晶体管的计算机模拟。例如，用 Early 电压介绍了 Early 效应，用电荷控制模型考虑了低发射结偏压和高发射结偏压效应。用相同的方法描述了基区渡越时间。从这一问题自然地引出了晶体管的电荷控制模型以及应用的实例。由电荷控制模型还推导出小信号晶体管(混合Π)模型，并且考虑了晶体管各种等效电路模型的关系。双极晶体管的频率响应表现出了同质和异质双极晶体管之间的差异。器件讨论集中于 pnp 型晶体管。

第8章　金属-氧化物-硅系统特性　本章提供了讨论 MOS 器件所需要的物理背景知识。描述了栅压控制的半导体表面积累、耗尽和反型状态，提出了 MOS 系统的平带电压和阈值电压的概念。选学内容为 pn 结的表面效应。本章最后讨论了用于存储电路中的 MOS 电容器和电荷耦合器件。

第9章　MOS场效应晶体管Ⅰ:物理效应和模型　首先利用近似的电荷控制模型推导了MOSFET的经典理论,然后计入了短沟道效应,迁移率降低,以及速度饱和效应。MOSFET尺寸的进一步缩小为先进的CMOS结构的快速发展提供了远景　本章还给出一个简化的电路设计模型和基于工业标准的BSIM模型的参数提取技术,为工艺发展和实用设计架起了桥梁。

第10章　MOS场效应晶体管Ⅱ:强场效应　本章介绍了MOSFET的基本描述中没有考虑的,但在现代CMOS工艺中非常重要的几个强场效应:沟道长度调制、输出电阻、衬底电流、栅极电流、MOSFET击穿、器件退化,以及MOS不挥发存储器结构.器件方面描述了利用强场效应的浮栅、不挥发存储单元的设计及存储单元与门阵列结构的集成。

目　录

第1章　半导体电学特性 ……………………………… (1)

1.1　半导体材料物理 ……………………………… (1)

1.1.1　固体的能带模型 ……………………………… (2)

1.1.2　空穴 ……………………………… (6)

1.1.3　成键模型 ……………………………… (8)

1.1.4　施主和受主 ……………………………… (9)

1.1.5　热平衡统计 ……………………………… (12)

1.2　半导体中的自由载流子 ……………………………… (21)

1.2.1　漂移速度 ……………………………… (23)

1.2.2　迁移率和散射 ……………………………… (24)

1.2.3　扩散电流 ……………………………… (30)

1.3　器件:霍尔效应磁传感器 ……………………………… (33)

1.3.1　霍尔效应的物理机制 ……………………………… (33)

1.3.2　集成霍尔效应磁传感器 ……………………………… (35)

小结 ……………………………… (37)

参考文献 ……………………………… (38)

参考书 ……………………………… (39)

习题 ……………………………… (39)

第2章　硅工艺 ……………………………… (48)

2.1　硅平面工艺 ……………………………… (48)

2.2　晶体生长 ……………………………… (54)

2.3　热氧化 ……………………………… (57)

氧化动力学 ……………………………… (58)

2.4　光刻和图形转移 ……………………………… (64)

2.5　掺杂和扩散 ……………………………… (69)

2.5.1　离子注入 ……………………………… (69)

2.5.2　扩散 ……………………………… (72)

2.6　化学气相淀积 ……………………………… (82)

2.6.1　外延 ……………………………… (82)

2.6.2　非外延薄膜 ……………………………… (83)

2.7　互连和封装 ……………………………… (90)

2.7.1　互连 ……………………………… (90)

2.7.2　测试和封装 ……………………………… (97)

2.7.3　污染 ……………………………… (98)

2.8　化合物半导体工艺 ……………………………… (98)

2.9　数值模拟 ……………………………… (101)

2.9.1　模拟的基本概念 ……………………………… (101)

2.9.2　网格 ……………………………… (102)

2.9.3　工艺模型 ……………………………… (103)

2.9.4 器件模拟 …………………………………………………………………… (109)
2.9.5 模拟面临的挑战 …………………………………………………………… (110)
2.10 器件:集成电路中的电阻 ……………………………………………………… (111)
小结 ……………………………………………………………………………………… (115)
参考文献 ………………………………………………………………………………… (116)
参考书 …………………………………………………………………………………… (117)
习题 ……………………………………………………………………………………… (117)
第3章 金属-半导体接触 …………………………………………………………… (121)
3.1 电子系统中的平衡 …………………………………………………………… (121)
3.2 理想的金属-半导体结 ………………………………………………………… (123)
3.2.1 能带图 ……………………………………………………………………… (123)
3.2.2 电荷、耗尽区和电容 ……………………………………………………… (126)
3.3 电流-电压特性 ………………………………………………………………… (132)
3.3.1 Schottky 势垒+ …………………………………………………………… (133)
3.3.2 Mott 势垒+ ………………………………………………………………… (136)
3.4 非整流(欧姆)接触 …………………………………………………………… (137)
3.4.1 隧道接触 …………………………………………………………………… (137)
3.4.2 Schottky 欧姆接触+ ……………………………………………………… (138)
3.5 表面效应 ……………………………………………………………………… (141)
3.5.1 表面态 ……………………………………………………………………… (141)
3.5.2 金属-半导体接触的表面效应+ …………………………………………… (142)
3.6 金属-半导体器件:Schottky 二极管 …………………………………………… (144)
小结 ……………………………………………………………………………………… (147)
参考文献 ………………………………………………………………………………… (149)
参考书 …………………………………………………………………………………… (149)
习题 ……………………………………………………………………………………… (149)
第4章 pn结 ………………………………………………………………………… (152)
4.1 缓变杂质分布 ………………………………………………………………… (152)
4.2 pn 结 …………………………………………………………………………… (158)
4.2.1 突变结 ……………………………………………………………………… (160)
4.2.2 线性缓变结 ………………………………………………………………… (166)
4.2.3 异质结 ……………………………………………………………………… (168)
4.3 反偏 pn 结 ……………………………………………………………………… (172)
4.4 结的击穿 ……………………………………………………………………… (176)
4.4.1 雪崩击穿+ ………………………………………………………………… (177)
4.4.2 Zener 击穿+ ………………………………………………………………… (182)
4.5 器件:结型场效应晶体管 ……………………………………………………… (184)
4.5.1 pn 结场效应晶体管(JFET) ………………………………………………… (184)
4.5.2 金属-半导体场效应晶体管(MESFET) ……………………………………… (190)
小结 ……………………………………………………………………………………… (192)
参考文献 ………………………………………………………………………………… (192)
参考书 …………………………………………………………………………………… (193)
习题 ……………………………………………………………………………………… (193)

第5章 pn结中的电流 ………………………………………… (197)
5.1 连续性方程 ………………………………………… (197)
5.2 产生与复合 ………………………………………… (199)
5.2.1 局域态:俘获和发射 ………………………………………… (199)
5.2.2 Shockley-Hall-Read 复合[+] ………………………………………… (201)
5.2.3 过剩载流子寿命 ………………………………………… (202)
5.3 pn结电流电压特性 ………………………………………… (206)
5.3.1 边界少数载流子浓度 ………………………………………… (207)
5.3.2 理想二极管分析 ………………………………………… (208)
5.3.3 空间电荷区电流[+] ………………………………………… (214)
5.3.4 异质结[+] ………………………………………… (217)
5.4 电荷存储与二极管瞬变特性 ………………………………………… (222)
5.5 器件建模和模拟 ………………………………………… (227)
5.5.1 集总元件模型 ………………………………………… (227)
5.5.2 分布式模拟[+] ………………………………………… (229)
5.6 器件 ………………………………………… (232)
5.6.1 集成电路二极管 ………………………………………… (232)
5.6.2 发光二极管 ………………………………………… (235)
小结 ………………………………………… (235)
参考文献 ………………………………………… (236)
参考书 ………………………………………… (236)
习题 ………………………………………… (237)
第6章 双极晶体管Ⅰ:基本特性 ………………………………………… (241)
6.1 晶体管工作原理 ………………………………………… (241)
6.1.1 原型晶体管 ………………………………………… (244)
6.1.2 集成电路晶体管 ………………………………………… (246)
6.2 放大偏置 ………………………………………… (247)
6.3 晶体管开关工作 ………………………………………… (256)
6.4 EBERS - MOLL 模型 ………………………………………… (260)
6.5 器件:平面双极放大和开关晶体管 ………………………………………… (264)
6.6 器件:异质结双极晶体管 ………………………………………… (271)
6.6.1 双异质结双极晶体管 ………………………………………… (274)
6.6.2 准中性基区的带隙缓变 ………………………………………… (275)
小结 ………………………………………… (277)
参考文献 ………………………………………… (278)
参考书 ………………………………………… (278)
习题 ………………………………………… (279)
第7章 双极晶体管Ⅱ:局限性与模型化 ………………………………………… (282)
7.1 基区宽度调变效应(Early 效应) ………………………………………… (282)
7.2 发射结低偏置效应和高偏置效应 ………………………………………… (285)
7.2.1 发射结低偏置电流 ………………………………………… (285)
7.2.2 大注入 ………………………………………… (286)

7.2.3 基区电阻 …… (291)
7.3 基区渡越时间 …… (295)
7.4 电荷控制模型 …… (298)
7.5 晶体管小信号模型 …… (308)
7.6 双极晶体管的频率限制 …… (313)
7.7 计算机模拟中的双极晶体管模型[*] …… (317)
7.8 器件:pnp 双极晶体管 …… (321)
7.8.1 衬底 pnp 晶体管 …… (321)
7.8.2 横向 pnp 晶体管 …… (321)
小结 …… (326)
参考文献 …… (326)
习题 …… (327)
第 8 章 金属-氧化物-半导体系统的性质 …… (332)
8.1 理想 MOS 结构 …… (332)
8.1.1 热平衡能带图 …… (333)
8.1.2 多晶硅栅和金属栅 …… (335)
8.1.3 平带电压 …… (336)
8.2 理想 MOS 结构的分析 …… (337)
8.3 MOS 电学特性 …… (340)
8.3.1 硅衬底中的电荷模型 …… (340)
8.3.2 热平衡状态 …… (340)
8.3.3 非平衡状态 …… (343)
8.4 MOS 系统的电容 …… (345)
8.4.1 理想 MOS 系统的 C-V 特性 …… (345)
8.4.2 C-V 测量中的实际因素 …… (348)
8.4.3 准静态(低频)C-V 测量 …… (349)
8.5 非理想 MOS 系统 …… (350)
8.5.1 氧化层及界面态电荷 …… (350)
8.5.2 氧化层电荷的起源 …… (353)
8.5.3 氧化层电荷的实验测定 …… (355)
8.6 pn 结的表面效应[†] …… (356)
8.7 MOS 电容器及电荷耦合器件 …… (359)
8.7.1 MOS 存储器 …… (360)
8.7.2 电荷耦合器件 …… (362)
小结 …… (365)
参考文献 …… (366)
习题 …… (366)
第 9 章 MOS 场效应晶体管 I:物理效应和模型 …… (371)
9.1 MOSFET 的基本工作原理 …… (373)
9.1.1 强反型区 …… (374)
9.1.2 沟道长度调制 …… (377)
9.1.3 衬偏效应 …… (378)
9.1.4 体电荷效应 …… (381)

9.1.5 离子注入调整阈值电压 …… (382)
9.1.6 耗尽型 MOSFET …… (384)
9.1.7 亚阈值区 …… (385)
9.1.8 小信号等效电路模型 …… (389)
9.2 短沟 MOSFET 的改进模型 …… (390)
9.2.1 长沟分析的局限性 …… (390)
9.2.2 短沟效应 …… (390)
9.2.3 迁移率下降 …… (395)
9.2.4 速度饱和 …… (397)
9.2.5 短沟 MOSFET 的漏极电流 …… (398)
9.2.6 MOSFET 的按比例缩小和短沟模型 …… (400)
9.3 器件:互补 MOSFET-CMOS …… (403)
9.3.1 CMOS 的设计考虑 …… (404)
9.3.2 MOSFET 参数和参数提取 …… (406)
9.3.3 CMOS 的闩锁效应+ …… (409)
9.4 展望未来 …… (413)
9.4.1 按比例缩小的目标 …… (413)
9.4.2 栅耦合 …… (413)
9.4.3 速度过冲 …… (415)
小结 …… (415)
参考文献 …… (417)
习题 …… (417)
第 10 章 MOS 场效应晶体管 II:强场效应 …… (423)
10.1 速度饱和区的电场 …… (423)
10.1.1 准二维模型 …… (424)
10.2 衬底电流 …… (429)
10.2.1 热载流子效应 …… (429)
10.2.2 衬底电流模型 …… (430)
10.2.3 衬底电流对漏极电流的影响 …… (434)
10.3 栅极电流 …… (435)
10.3.1 幸运电子模型 …… (435)
10.3.2 低栅压载流子注入 …… (439)
10.3.3 p 沟道 MOSFET 的栅极电流 …… (440)
10.4 器件退化 …… (440)
10.4.1 n 沟道 MOSFET 的退化机制 …… (440)
10.4.2 n 沟道 MOSFET 退化的表征 …… (441)
10.4.3 器件寿命的加速测量 …… (442)
10.4.4 减小漏场的结构 …… (443)
10.4.5 p 沟道 MOSFET 的退化 …… (445)
10.5 器件:MOS 不挥发存储器结构 …… (445)
10.5.1 浮栅存储单元的编程 …… (447)
10.5.2 浮栅存储单元的擦除 …… (449)
10.5.3 浮栅存储阵列 …… (450)

小结 …………………………………………………………………………………… (451)
参考文献 ……………………………………………………………………………… (452)
习题 …………………………………………………………………………………… (453)
部分参考答案 ………………………………………………………………………… (455)
附录 ………………………………………………………………………………… (457)
索引 ………………………………………………………………………………… (465)

第1章　半导体电学特性

从日常的生活经验中，我们知道材料的电学特性会在很大的范围内变化。当电压 V 加在一块横截面积均匀的棒状样品的两端时，测量流过的电流 I 便可以计算出其电阻 $R = V/I$。电阻率（样品材料的基本电学特性之一）与棒状样品的电阻和几何尺寸之间有如下关系

$$\rho = R\frac{A}{L} \tag{1.1}$$

式中 L 和 A 分别是样品的长度和横截面积。

固体器件中常用材料的电阻率可在很宽的范围内变化。以制作典型硅集成电路时所用的各种材料的室温电阻率为例，连接集成电路各元件的最常用的金属是铝和铜，它们的电阻率非常低，约为 $10^{-6}\Omega\cdot\text{cm}$。隔离集成电路各部分所用的材料为绝缘体，它们的电阻率则分布在另一个极端。例如二氧化硅，电阻率大约为 $10^{16}\Omega\cdot\text{cm}$，比铝的大 22 个数量级。用于集成电路封装的塑料的电阻率则高达 $10^{18}\Omega\cdot\text{cm}$。这样，典型的集成电路中的材料电阻率变化范围可达 24 个数量级。

在电学上，我们通常根据材料的电阻率对它们进行分类。电阻率小于 $10^{-2}\Omega\cdot\text{cm}$ 的材料称为导体，电阻率大于 $10^{5}\Omega\cdot\text{cm}$ 的材料称为绝缘体，而电阻率位于两者之间的一类材料称为半导体。半导体有一个非常重要的电学特性，即电阻率可以人为改变且精确控制；同时，可以根据需要制备成两种导电类型中的任一种。本书的主要内容就是如何应用半导体的这些独特的性质

本章第一部分先解释固体的电阻率变化范围如此之大的原因。然后回顾半导体物理的一些重要概念。由于本书主要针对器件，我们对材料的讨论相对比较简单，而且假定读者已具有参考文献[1]中的基本半导体物理知识和固体物理知识，这里只回顾对后面学习有用的一些重要概念。本章附录还简要回顾了电磁理论中的一些概念。

1.1　半导体材料物理

研究孤立原子中的电子能够帮助我们了解固体中电子的物理特性，而研究电子在固体中受到的微扰，就可以得到固体的一些重要的电学特性。

下面先考虑电子只受到孤立原子的原子核作用时所具有的能级。然后，考虑当其他原子靠近这个原子，并且其他原子的原子核也会对电子有作用时的情况，这正是固体中的情形。通过这种方式可以研究晶体中的原子核对相关电子的行为的影响。最后，我们研究外加电场对固体中的电子的作用。

半导体物理中要用到两个互补的模型：能带模型和晶体成键模型。我们先介绍能带模型，然后引入晶体成键模型并介绍其他相关概念。

1.1.1 固体的能带模型

我们知道,原子核对电子有库仑作用,电子因而具有某些特定的允许能量(如图1.1)。相对于参考能量 $E=0$,电子只能占据下面一系列能级中的某一个

$$E_n = \frac{-Z^2 m_0 q^4}{8\epsilon_0^2 h^2 n^2} \tag{1.1.1}$$

式中电子具有 $-q$ 电荷,所以(1.1.1)式中的电荷 q 为正。为方便起见,本书中 q 都取正值。公式(1.1.1)中的其他符号的含义为:Z——原子核中的质子数;m_0——自由电子质量;ϵ_0——真空电容率[1];h——普朗克常数;n——任意正整数。对氢原子,$Z=1$,其电子的允许能级低于零参考能级,等于 $-2.19\times10^{-18}/n^2$ J 或 $-13.6/n^2$ 电子伏特(eV)[2]。低温下,当原子中不止有一个电子时,电子将从能量最低的允许能级开始填充。根据Pauli不相容原理,每一能级最多可容纳两个电子(自旋相反)。

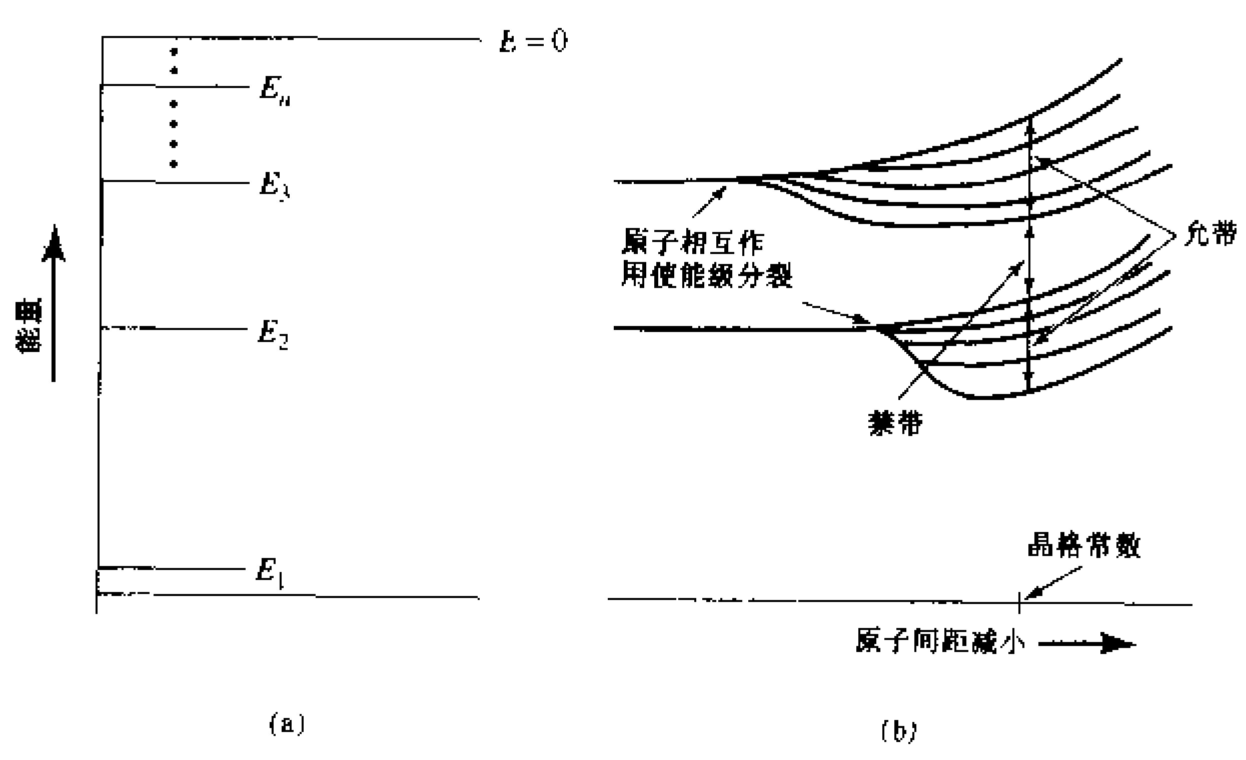


图1.1 (a)电子受到原子核的库仑作用形成分裂的允许能级;(b)当原子间的距离减小时,分立的电子能态分裂为允带,中间被禁带分开;晶体材料的电学特性与允带能量和禁带宽度有关,而这些能量又决定于原子间距及晶格常数

我们考虑占据在最高能级上的电子(不考虑原子中低填充能级)的情况。当两个原子之间的距离较大时,每个原子的电子能量 E_n 由公式(1.1.1)给出。而当两个原子彼此靠近时,一个原子的原子核会对另一个原子中的电子有作用,因而改变了电子的势能,由于这一势能的变化,所有电子的允许能级都被改变。

描述两原子系统时,必须考虑Pauli不相容原理。当系统由彼此孤立的两个原子组成时,它们之间的相互作用可以忽略,各自的 E_n 能级上最多只能容纳两个自旋相反的电子,所以系

1 有时也称做介电常数。

2 在半导体物理中采用电子伏特单位($1\text{eV}=1.6\times10^{-19}$J,参见封面内页的转换关系)非常方便。尽管电子伏特不是MKS制单位,但用电子伏特表示能量能避免复杂的表达式,因而被广泛应用,本书也采用电子伏特单位。

统中最多可包含四个电子。但是，当两个原子彼此靠近有相互作用时，由于一个能级上只能有两个电子，孤立时彼此相同的电子能级就要解除简并，分裂成两个能量略微不同的能级，这时系统才能容纳四个电子。因此，当两个原子相互靠近时，孤立原子中的每一E_n能级受到扰动，从而分裂成能量相隔很小的两个能级。如果原子进一步靠近，则相互作用更强，能级分裂将进一步增加。

当大量原子结合在一起构成晶体时，每个电子受到的库仑作用力进一步改变，电子能级的变化也更剧烈。同样，Pauli 不相容原理要求每一允许的电子能级能量彼此略有不同，所以晶体中将有很多紧密排列的不同能级。孤立原子相应的每一分立能级都分裂成许多能级，这样形成的每一组能级对应于孤立原子中的一个能级。当系统中含有 N 个原子时，原来的能级 E_n 分裂成 N 个不同的允许能级，构成一个能带，最多可容纳 $2N$ 个电子(考虑电子自旋简并)。孤立原子的两个允许能级的分裂如图 1.1(b)所示，图中给出了原子间距连续变化时允许能级的变化规律。在这种情况下，孤立原子的分立能级在原子间距减小时分立成许多能级。当原子间距等于晶格常数时(如图 1.1(b)所示)，两个允许能带通常隔着一个不能有电子占据的禁带。机械和电学谐振子也存在与原子能级分裂相似的情况，它们具有分立的谐振频率，当许多相似的谐振子系统相互作用时则会存在多个谐振频率。

由于晶体中的原子数可达 $10^{22}\,cm^{-3}$量级，而总的能带宽度仅为几个电子伏特，因此每一个能带中 N 个不同能级的间距远小于电子在室温下的热能(~1/40eV)，从而电子很容易在各能级间移动。这样我们可以说包含有 $2N$ 个电子的允许能带中，能级是连续分布的。这一允许带与相邻允许带之间或者被禁带隔开，如图 1.2(a)所示，或者与相邻能带重叠。给定材料能带的具体特征(能带间是相互重叠还是形成禁带，如果形成禁带，禁带宽度是多少?)基本上决定了材料的电学特性。这也是区分导体、绝缘体和半导体的本质特征。

尽管孤立原子的每一能级都分裂成了由 $2N$ 个能级组成的能带，但每一允许能带的宽度不一样。能量高的能带比能量低的能带宽。考虑与第 n 个能级相关的波尔半径 r_n 可以理解这一差异

$$r_n = \frac{n^2\epsilon_0 h^2}{Z\pi m_0 q^2} = \frac{n^2}{Z} \times 0.0529\ \text{nm} \tag{1.1.2}$$

对较高能级上(对应的 n 大)的电子，由于受到原子核的束缚力弱且远离原子核，这就与相邻原子更接近，因而受到相邻原子的影响更强烈。这一强烈的相互作用引起能级的变化也更大，导致与孤立原子中高能级对应的能带将更宽。

晶体处于平衡态时的一维能带图如图 1.2(a)所示。每一能带中最高的允许能级与上一个能带中最低的允许能级间由带隙 E_g 隔开。很容易将这个能带图扩展为二维(图 1.2(b))，垂直方向仍表示电子能量，水平方向表示半导体晶体中的位置。这种表示方法强调了能带中的电子不与任一单独的原子核相联系，只受到晶体边界的限制。在考虑不同能带结构的材料组成的半导体器件时，这种能带图特别有用。我们在简单介绍晶体能带理论的基础时，明确假定每一原子与其相邻原子的各种特性包括取向都相同。换句话讲，考虑的是完美晶体。事实上，满足器件质量要求的半导体材料通常具有极好的晶格有序性，其中缺陷只有十亿分之几甚至更少。

任何元素的原子结合形成固体时，分立能级都会形成能带。然而，不同的固体能带中电子

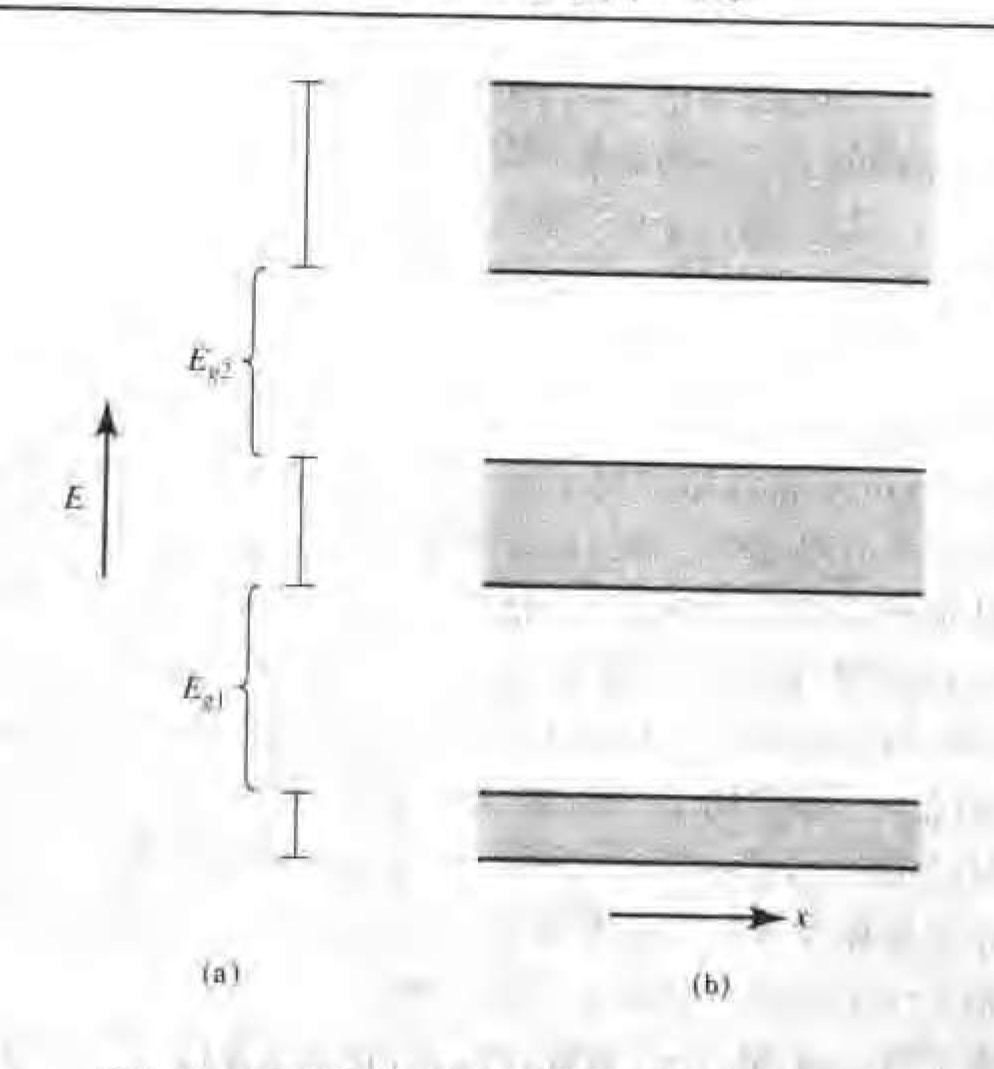


图 1.2　当固体中较多的原子影响到每一个电子时，允许能级展宽形成被禁带隔开的允带。(a)一维表示法；(b)能量随位置变化的两维能带图

数目的不同，强烈地影响到其电学特性。例如，考虑由 N 个原子组成的碱金属，原子的外壳层只有一个价电子。当这些原子彼此靠近时，这一能级形成能带，在最简单的情况下，这一能带可容纳 $2N$ 个电子。N 个价电子填充了能带的下半部分(见图 1.3(a))，在填充态上面还有空的能态。靠近填充态顶部的电子便很容易从外加电场获得少量的能量进入这些空状态中去，这些电子几乎像是自由电子，在外加电场的作用下可以在晶体中移动。总之，金属具有部分填充的能带，因而具有高的导电性。

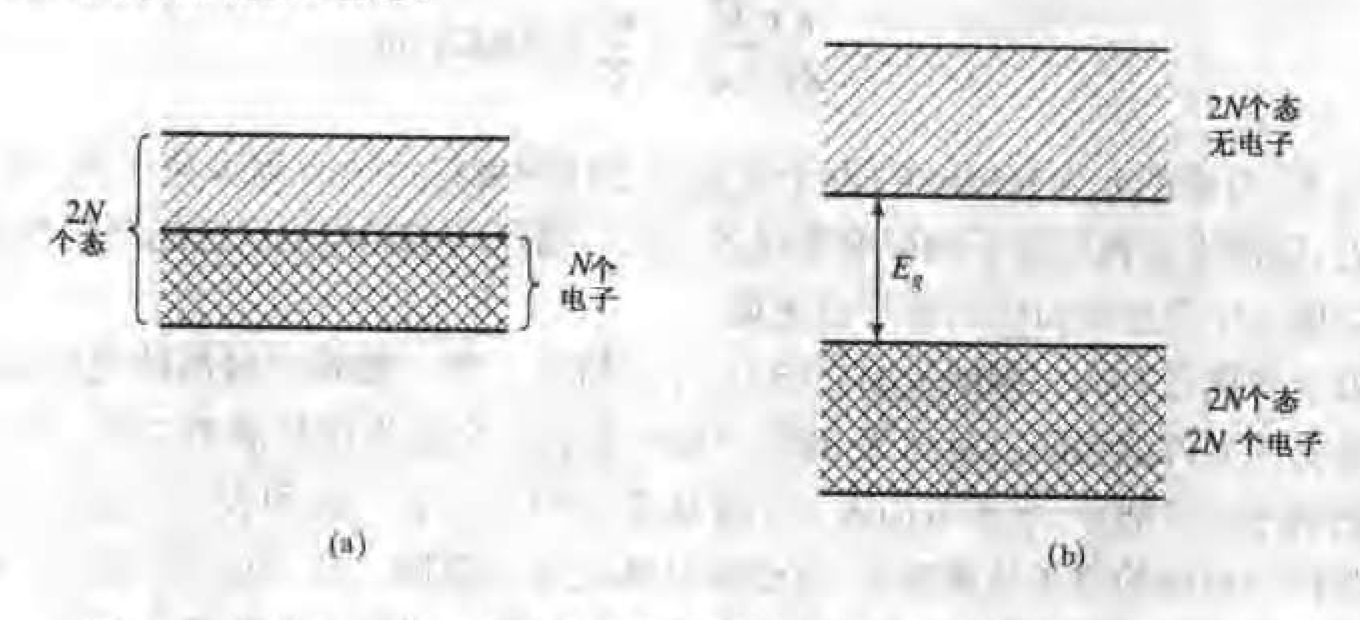


图 1.3　能带图。(a) N 个电子填充了 $2N$ 个允许态中的一半，这是金属中的情形；(b)一个全空的能带与被 $2N$ 个电子全部填满的能带之间由带隙 E_g 分开，这是绝缘体的情况

与上述电学特性明显不同的情况是，材料中的价带(最外壳层)完全填满了电子，并与上

面更高的能带之间存在带隙。这种情况下，与填充能带最近的允许带在低温下是全空的，如图1.3(b)所示，这正是绝缘体的特征。空带中能量最低的能级与填充带中最高能级间的能量差为带隙 E_g。绝缘体的 E_g 一般约为5eV（SiO_2 为8~9eV）[③]，远大于典型的热运动能量和电场提供给电子的能量（十分之几个eV或更小）。理想情况下，空的允许态附近没有电子存在，因此也没有可从外电场中获得能量的电子，从而也就没有电子输运而产生的电流了。这样的材料为绝缘体。

为帮助理解问题，下面采用一个比喻。用一个两端封闭水平放置的玻璃试管表示允许能态，试管中的液体表示固体中的电子。与金属对应的情况是试管内部分充有液体（图1.4）。当倾斜试管以施加外力（这里是重力）时，液体很容易沿试管流动。与绝缘体对应的情况是试管完全被液体充满（图1.5）。当试管倾斜时，因为试管中没有让液体流动的空间（即没有空的允许态），则无液体流动。

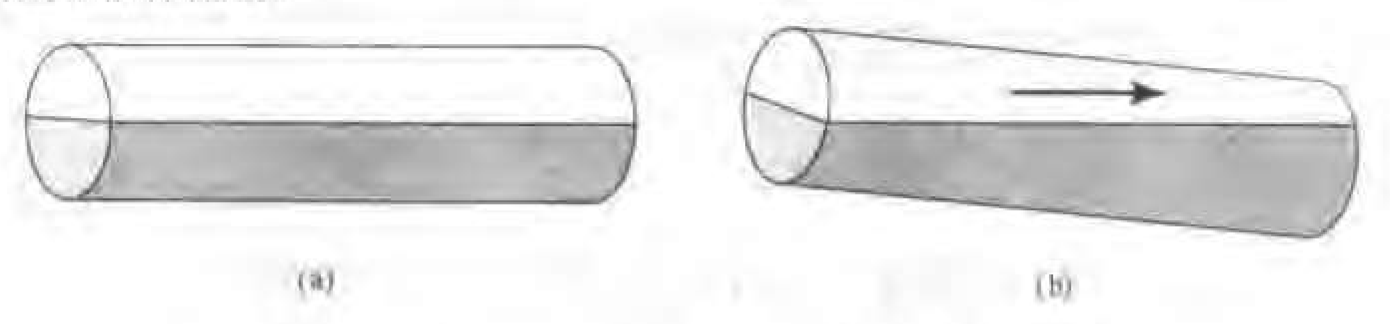


图1.4 电子在允带中的运动类似于两端封闭的玻璃试管中液体的流动；在半满的试管中液体可以流动，正如电子在金属中的运动一样

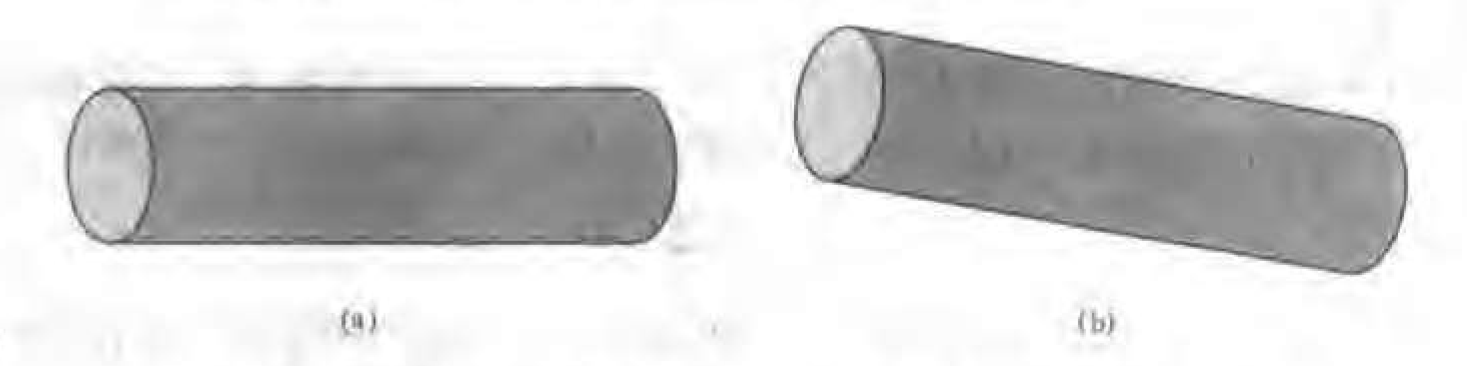


图1.5 在两端封闭的完全充满液体的试管中不会出现液体流动现象

绝缘体和半导体具有相似的能带结构，它们之间电学性能的差异来自禁带宽度的不同。另外，半导体中还可掺入能增加导电性的杂质以增强近乎空的能带的导电能力。在绝对零度下，半导体中最高填充能带与最低空能带间带隙的典型值为1eV量级（硅：1.1eV；锗：0.7eV；砷化镓：1.4eV）。在不含杂质的半导体中，最高的填充能带被孤立原子中的价电子占据，因此这个能带称为价带。

半导体的能带结构如图1.6所示。高于绝对零度时，因为少数电子拥有足够的热能可被激发越过禁带、进入上面的允许能带中，因而价带没有被完全填满。带隙越小，温度越高，能在两个能带之间跃迁的电子数目越多。在高能带上的电子可以很容易获得少量的能量形成受外加电场控制的电流。因为进入此空能带的电子可以参与导电，这个能带被称为导带。导带中的电流密度 J（单位面积电流）等于导带中电子电荷（$-q$）乘以每个电子的速度（v_i），然后对材

③ 如何确定非晶材料的带隙引起了理论学家的困惑，但是非晶材料短程有序，也可以借助于上面对能带的描述来解释其特性。

料单位体积内所有导带电子求和。

$$J_{cb} = \frac{I}{A} = \sum_{cb}(-q)v_i \tag{1.1.3}$$

因为导带中只有少数电子,所以在给定电场下的电流将远小于金属。

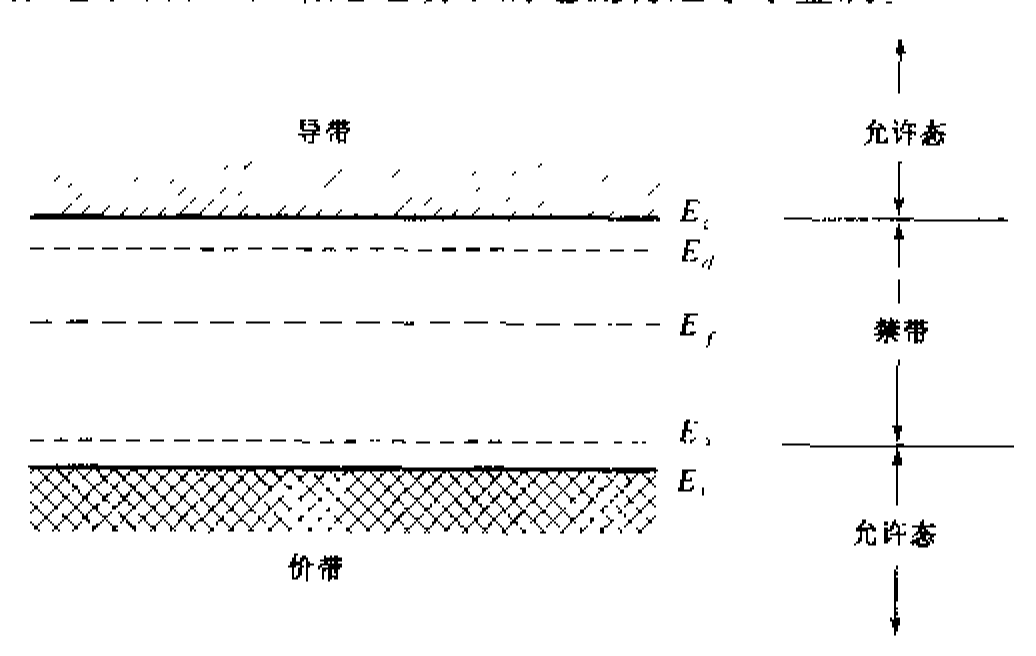


图 1.6 半导体的能带图：E_c 为导带边，E_d 为禁带中的施主能级，E_f 为 Fermi 能级，E_a 为受主能级，E_v 为价带边

1.1.2 空穴

当电子被激发到导带时,价带中就留下了空态。如果存在外加电场,电子可在电场的作用下进入这些空态而产生电流。单位体积下对价带中所有电子的运动求和便可得到这个电流

$$J_{vb} = \sum_{vb}(-q)v_i \tag{1.1.4}$$

因为价带中有很多电子,而只有很少的空态,因此描述电子与这些空态的相互作用对导电性能的影响比描述价带中所有电子的运动要容易得多。在数学上,我们可以把价带电流表述为价带完全填满时对应的电流减去与失去的电子相对应的电流之差。这样,对单位体积内的全部电子求和可得

$$J_{vb} = \sum_{vb}(-q)v_i = \sum_{满带}(-q)v_i - \sum_{空态}(-q)v_i \tag{1.1.5}$$

由于在完全填满的能带中没有电流流过(因为占据在能带上的电子不能获得净的能量),价带电流可写为

$$J_{vb} = 0 - \sum_{空态}(-q)v_i = \sum_{空态} qv_i \tag{1.1.6}$$

式中是对单位体积内所有的空态进行求和。等式 1.1.6 表明把空态看成带正电的粒子后,便可用空态来表示价带中的电荷运动了。这些"粒子"被称做空穴。只有在讨论固体能带时才能用空穴的概念,在自由空间中不能使用。注意图 1.3 和图 1.6 都是电子的能带图,当电子向能带图上部移动时电子能量增加。然而,由于电荷符号相反,在同一能带图中空穴的能量向下

是增加的[4]。

空穴的概念可继续用玻璃试管中的液体进行类比。考虑两个密封的玻璃试管，一个完全充满，一个是空的（图1.7a）。当试管倾斜时，没有液体的流动发生（图1.7b）。如果将少量的液体从下面的试管转移到上面的试管中（图1.7c），这时当上面的试管倾斜时，液体就可以流动了（图1.7d）。这种流动好比是固体导带中电子的传导。下面的试管由于移去一部分液体出现了空隙，这一空隙好比是价带中的空穴。它不能存在于近乎充满的试管的外部，就像只有在讨论近乎填满的价带时，空穴的概念才有用一样。当试管倾斜时，液体向下流动，而管中的空隙向相反的方向运动，好像具有与液体相反符号的质量一样。同样，价带中的空穴与电子的运动方向相反，好像具有与电子相反的电荷一样。正如讨论试管中少量空隙的运动比讨论大量液体的运动容易，讨论少量空穴的运动比讨论几乎填满价带的大量电子的运动要容易得多。

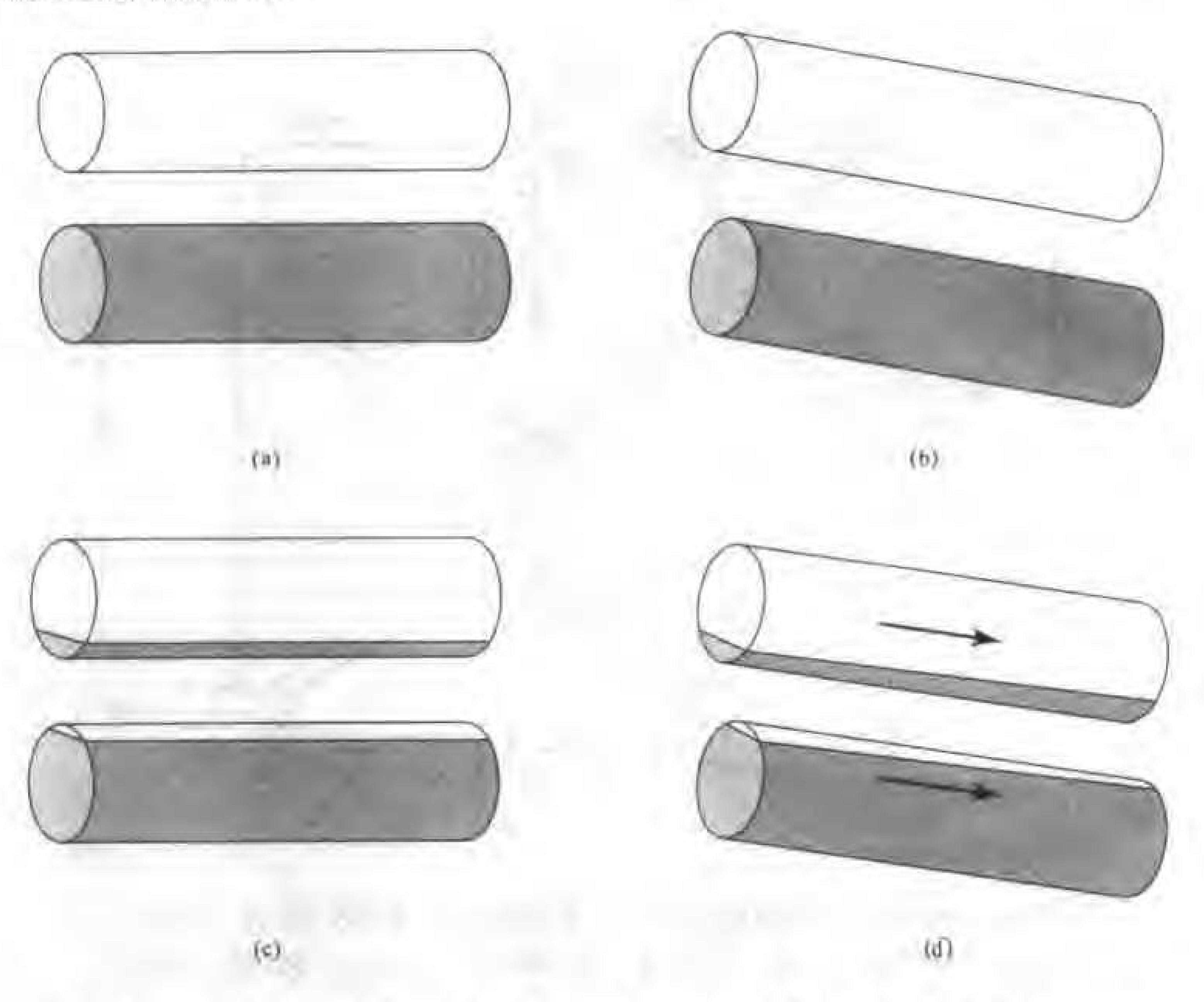


图1.7　用试管中的液体类比半导体。(a)和(b)在充满的和空的试管中没有发生液体的流动　(c)和(d)如果将部分液体从充满的试管转移到空试管中，下面的试管会留有空隙，这时液体在两个试管中都可以流动

④　如图1.6的能带图经常被简化，只画出价带边（由 E_v 表示）和导带边（由 E_c 表示），因为我们主要对靠近这两个能级的状态和将它们分隔开的带隙 E_g 感兴趣。

1.1.3　成键模型

对半导体中自由空穴和自由电子的讨论也可利用晶体中饱和的电子键和断裂的电子键的概念。这一观点通常称为成键模型，它虽然不能解释重要的量子力学对晶体中电子行为的限制，但是却能定性地说明一些有用的概念。

为讨论成键模型，我们考虑如图 1.8 所示的金刚石结构，硅和锗都具有这样的晶体结构。金刚石结构中，每一原子与它最近邻的四个原子形成共价键。每个键由两个紧密结合在一起的电子组成，其中每一个原子贡献一个电子。绝对零度时，所有的电子被束缚在这些键中，电子不能在外电场的作用下在晶体中自由运动。当温度升高时，热能会打破某些键产生出几乎是自由的电子，在外电场的作用下可对电流有贡献。这一电流对应于能带模型中的与导带相关的电流。

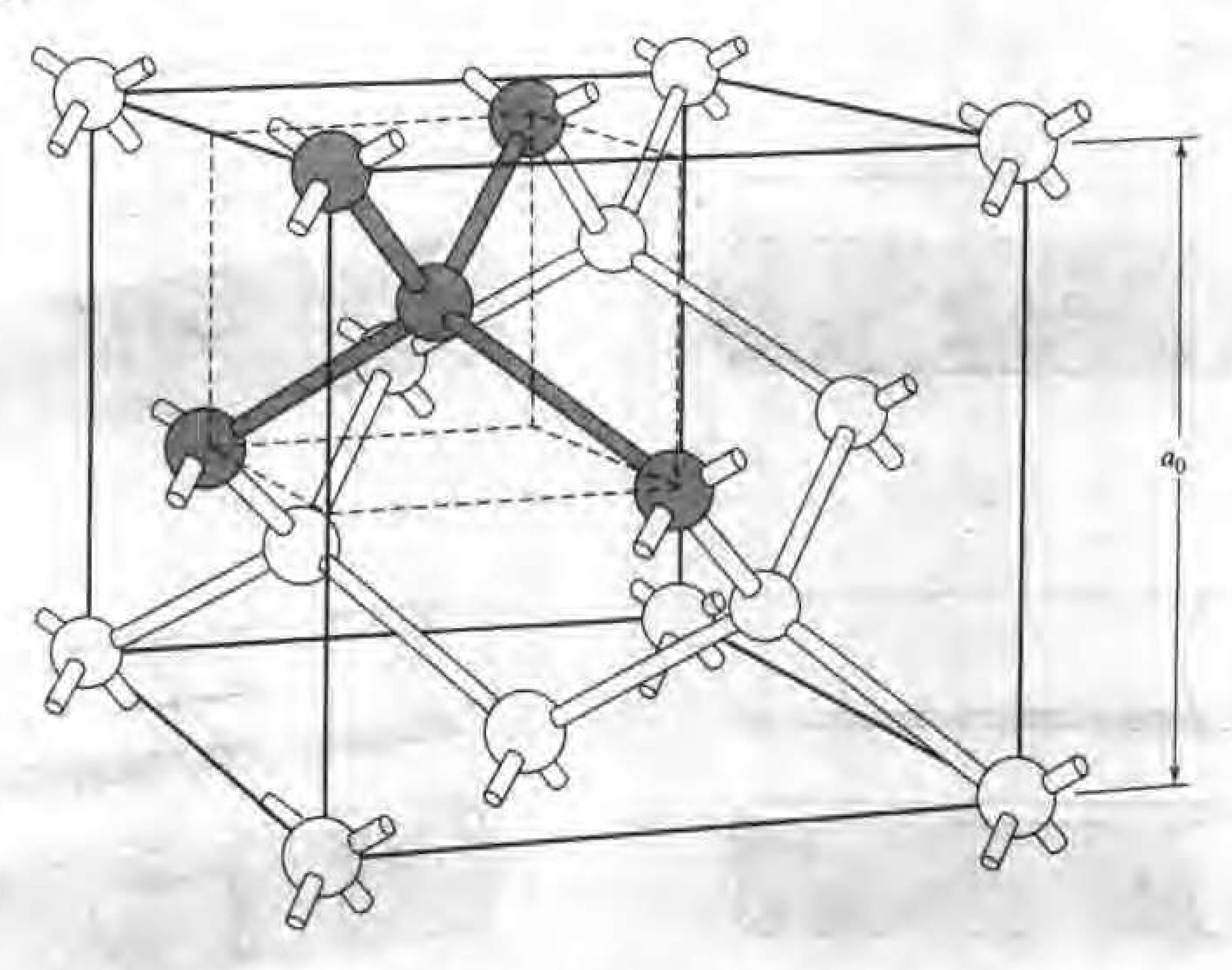


图 1.8　金刚石结构的特征是每个原子与近邻的四个原子形成共价键。金刚石、硅和锗的晶格常数用 a_0 表示，分别等于 0.356nm，0.543nm 和 0.565nm。最近邻原子间距为($\sqrt{3}a_0/4$)。图中的 18 个原子中只有 8 个属于体积 a_0^3 的。其中顶角上的原子被 8 个同样的立方体共用，它们总的贡献为 1 个原子；六个面上的原子被两个同样的立方体共用，总贡献为 3 个原子；还有 4 个原子位于立方体内。因此原子密度为 $8/a_0^3$，分别等于 17.7，5.00 和 $4.43\times10^{22}\,cm^{-3}$。（引自 W. Shockley：半导体中的电子和空穴，Van Nostrand，Princeton，N. J.，1950）

当一个键被热能打破后，自由了的电子离开原子，留下一个空键。然后，相邻键的电子可以跳到这个空键上，又留下另一个空的键。因此空键与电子的移动方向相反。如果电子在电

场作用下运动，空键则沿与电子相反的方向运动，好像带有正电荷[5]。这种空键对应于能带图中价带内的空穴。

闪锌矿结构与金刚石结构有很多相似之处，几种重要的由周期表中第三族和第五族元素组成的化合物半导体（称为Ⅲ-Ⅴ族半导体）就属于这种结构。某些Ⅲ-Ⅴ族半导体，尤其是砷化镓（GaAs）和磷化镓（GaP）都是制造半导体器件的主要材料。本章末的表1.3中给出了元素半导体和化合物半导体许多重要的性能参数。表1.3中还包括了集成电路中用到的一些绝缘材料的相关数据。在本章末的第二个表（表1.4）中给出了最重要的半导体材料——硅的更多的性能。

1.1.4　施主和受主

迄今为止，我们一直讨论的是纯净的半导体材料，每个被激发到导带的电子会在价带中留下一个空态。因此，导带中带负电荷的电子浓度 n 等于价带中带正电荷的空穴浓度 p。这样的材料被称为本征半导体，电子和空穴浓度通常用下标 i（即 n_i 和 p_i）表示。然而，半导体最重要的应用常要涉及相邻的半导体材料间的相互作用，这些材料通常具有浓度不同的两类载流子。可以在空间上将两种不同带隙的材料结合在一起（如4.2节），或改变一种半导体材料中的载流子数目（这里讨论）来获得这样的结构。

在半导体中控制载流子数目最有效的方法是引入替位式杂质，即杂质占据纯净半导体中原子的位置。如果用周期表中Ⅴ族元素的一个杂质原子，比如磷（五个价电子），替代一个硅原子（四个价电子），杂质原子中的四个价电子在杂质与相邻硅原子间成键。第五个电子没有与相邻原子形成共价键，它只受到杂质原子中原子核的过剩正电荷弱的束缚作用。这时只需要很小的能量就可以破坏这一弱的束缚，所以第五个电子可以在晶体中移动，从而对电导有贡献。由于替位式的Ⅴ族杂质可给硅导带提供电子，所以称为施主。

为了估算打破施主原子对电子的束缚所需要的能量，需要考虑电子受到这个原子核的净的库仑势。假设电子只受到杂质原子核一个净的正电荷的作用，且这个作用因周围硅原子中其他电荷的极化效应而减弱。电子受到施主原子核的束缚能为

$$E=\frac{m_n^* q^4}{8h^2\epsilon_0^2\epsilon_r^2}=\frac{13.6}{\epsilon_r^2}\frac{m_n^*}{m_0}\text{eV} \tag{1.1.7}$$

式中，ϵ_r 为半导体的相对电容率；m_n^* 为半导体导带电子的有效质量。有效质量反映了晶格对电子运动的影响。硅的 $\epsilon_r=11.7$，$m_n^*=0.26m_0$，$E\approx0.03\text{eV}$，大约只有硅带隙能量（1.1eV）的3%（更详细的计算和测量表明典型的施主的束缚能会略微高一些，例如，磷为0.044eV，砷为0.049eV，锑为0.039eV）。由于束缚能很小，使得打破施主原子与第五个电子间弱的束缚比打破硅—硅共价键要容易得多。

n型半导体　根据能带模型，只需要很少的能量就可以将施主原子中的电子激发到导带，而将电子从价带激发到导带所需的能量则要大得多。因此可以将束缚在施主原子上的电子所对应的能态用导带边 E_c 下面大约0.05eV处的能级 E_d 表示（图1.6）。施主浓度（原子数/cm^3）通常用 N_d 表示。温度高于150K时，热能就足以将施主原子中的电子激发到导带。一旦电子

[5] 正电荷与空键有关，因为核外电子电荷不足以抵消原子核中质子的正电荷。

被激发到导带,晶格中留下一个固定的带正电荷的离子实。因此,施主提供的允许能态(施主能级)在有电子占据时为中性,空的时候带正电[⑥]。

如果大部分杂质是施主型的,导带中电子数目将远大于价带中空穴数目,这时称电子为多数载流子(或多子),称空穴为少数载流子(或少子)。该材料称为 n 型半导体,这时大部分电流是由带负电荷的电子传输的。图 1.9 给出了硅和锗中导带电子浓度随温度的变化关系。这张图清楚地表明,在足以离化施主原子(约 150K)而还不足以从硅—硅键中释放出电子(约 600K)的温度范围内,由于空穴浓度最多等于 n_i,电子浓度将远高于空穴浓度。

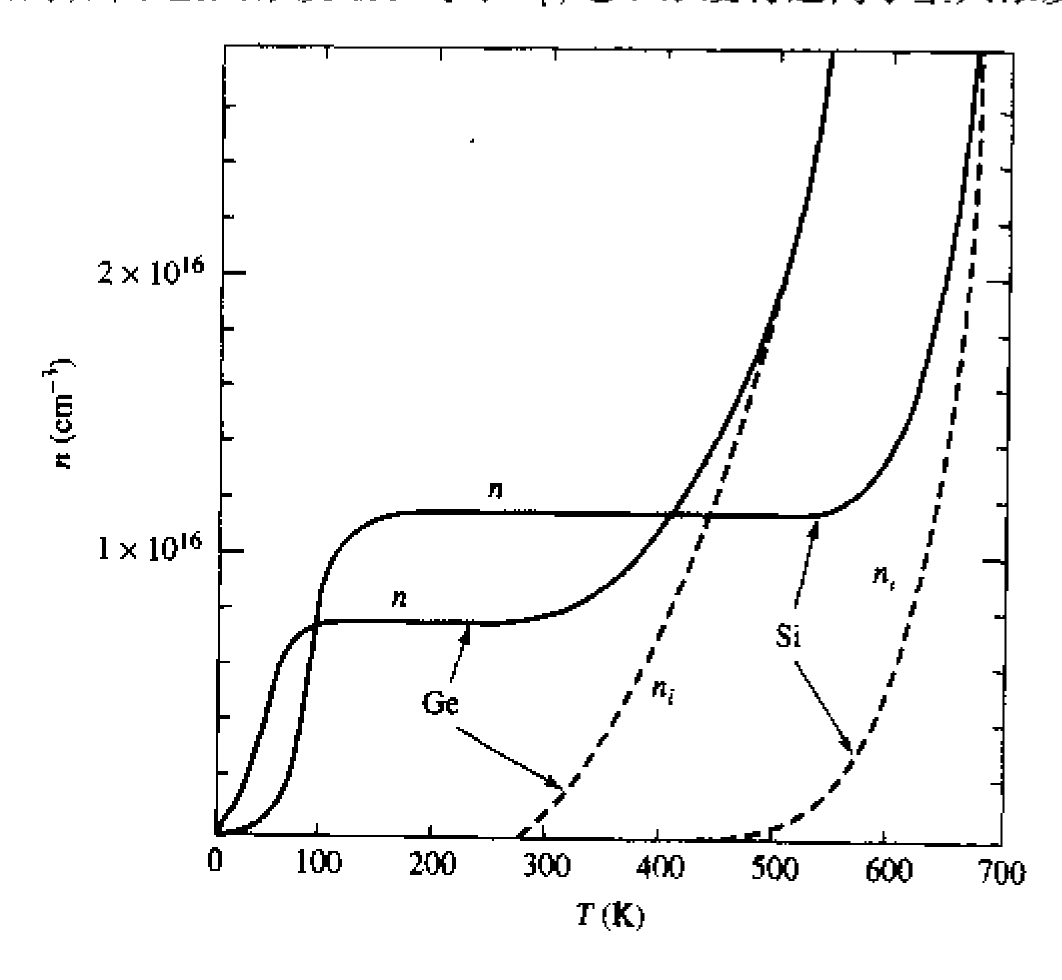


图 1.9 两种 n 型半导体材料的电子浓度随温度的变化

(a) 硅中 As 原子的掺杂浓度为 $1.15\times10^{16}\text{cm}^{-3}$[1]

(b) 锗中 As 原子的掺杂浓度为 $7.5\times10^{15}\text{cm}^{-3}$[2]

p 型半导体 与前面类似,具有三个价电子的杂质原子,如硼,也能在晶格中替代硅原子的位置。这三个电子与硅中四个价电子中的三个形成共价键,留下一个空键。如果附近键上的电子运动过来填入这个空键,则相当于空键带着正电荷运动,形成了空穴电导。正如对施主原子只需要少量能量就可以将电子激发到导带一样,从价带激发电子到由三价杂质引起的空键也只需少量的能量。这一能量用略高于价带边 E_v 的能级 E_a 来表示(图 1.6)。因为形成了容易接受电子的空键而对空穴导电有贡献的杂质称为受主杂质。受主浓度(原子数/cm^3)用 N_a 表示。如果固体中大部分杂质是受主,电流则主要来自带正电荷的空穴的传导,该材料称为 p 型半导体。受主能级空时是中性的,被电子占据时带负电。

半导体的导电性如果主要来自杂质原子的贡献,就称其为非本征半导体。施主和受主杂质如果是为改变载流子浓度而有意引入的,则称为掺杂原子。

在化合物半导体中,例如 GaAs,某些Ⅳ族杂质可替代其中任一种元素的晶格位置。这样,

⑥ 在本研究领域之外,例如化学中,对施主的定义不同,可能会引起某些混淆。

引入 GaAs 中的杂质硅如果替代 As 位，则可提供价带空穴，替代 Ga 位则可提供导带电子。这种双性的掺杂行为很难控制，但并不是所有Ⅳ族杂质都这样。例如，Ⅳ族元素 Sn 在 GaAs 中几乎毫无例外地替代 Ga 位，是有效的 n 型掺杂剂。来自Ⅵ族的杂质将替代 As 位，如 Te、Se 或 S，也用做 GaAs 的 n 型掺杂剂。Ⅱ族元素如 Zn 或 Cd 则一直被用做 p 型掺杂剂。

其他杂质原子或晶体缺陷可能提供对电子束缚较紧的能态，因而需要较大能量才能将电子从束缚态激发到导带。与浅施主能级离导带边能量差只有热能 kT 的几倍的情况不同，这样的深施主态不能用导带边以下较低的能级来表示。同样，深受主能级位于价带边以上较高的位置。由于深能级与杂质原子间的关系不像浅施主和浅受主那样直观，究竟属于施主还是受主要根据能级可能的带电状态来区分。深施主态指有电子占据时是中性的，无电子时带正电。而深受主指无电子时是中性的，被电子占据时带负电。

补偿　用浅施主杂质掺杂可得到 n 型材料，用浅受主杂质掺杂可得到 p 型材料，这是制备硅器件最重要的工艺步骤。掺杂工艺还有一种非常有用的方法，即可在后续工艺加入相反类型的掺杂剂（这里是 p 型杂质）来补偿已掺杂的硅晶体（例如 n 型样品）。参考图 1.6 可以帮助阐明这一过程。在图 1.6 中，施主原子在导带边 E_c 附近引入了允许能级 E_d，而受主原子在价带边 E_v 附近引入了允许能级 E_a。在典型硅器件的工作温度下，每一施主原子会失去一个电子而每一受主原子将得到一个电子。因为受主原子所处的能量状态低于施主能级，只要受主能级上有空态，施主能级上的电子就会转移（或向下跃迁）到更低能态的受主能级。因此在掺杂半导体中，有效掺杂浓度等于施主和受主浓度之差 $|N_d - N_a|$。如果 N_d 大于 N_a 则半导体为 n 型，反之为 p 型。虽然从理论上讲，可以通过补偿（$N_d = N_a$）得到零有效掺杂浓度，但这在技术上是不切实际的。第 2 章中（主要讨论工艺）我们将会看到，补偿的掺杂浓度通常必须要比原来的杂质浓度高出一个数量级。

例题　施主和受主

已知硅晶体中含有原子比为 10^{-6} 的 As 杂质，然后再均匀掺入 $3\times10^{16}\,\text{cm}^{-3}$ 的磷（P）原子和 $10^{18}\,\text{cm}^{-3}$ 的硼（B）原子，假设热退火后所有杂质完全激活。请问：

（a）硅样品的导电类型；

（b）多数载流子浓度。

解：As 是Ⅴ族杂质，为施主。由于硅的原子密度为 $5\times10^{22}\,\text{cm}^{-3}$（见表 1.3），原子比等于 10^{-6}，表明 As 的掺杂浓度为

$$5\times10^{22}\times10^{-6} = 5\times10^{16}\ \text{As原子数cm}^{-3}$$

再掺入 $3\times10^{16}\,\text{cm}^{-3}$ 的 P 原子，使得晶体中的施主掺杂浓度达到 $8\times10^{16}\,\text{cm}^{-3}$。

另外掺入杂质 B（Ⅲ族杂质）将硅的导电类型从 n 型转为 p 型，因为现在的受主浓度超过施主浓度。由于施主补偿，净的受主浓度低于 B 原子浓度。

（a）因此，硅样品为 p 型。

（b）空穴浓度等于净的掺杂浓度

$$\begin{aligned} p &= N_a(\text{B}) - [N_d(\text{As}) + N_d(\text{P})] \\ &= 10^{18} - [5\times10^{16} + 3\times10^{16}] \\ &= 9.2\times10^{17}\,\text{cm}^{-3} \end{aligned}$$

1.1.5 热平衡统计

在对半导体的导电特性进行更详细的讨论之前,我们先介绍另外三个概念:第一,热平衡的概念;第二,半导体处于热平衡时多数载流子和少数载流子浓度之间的关系;第三,采用 Fermi 统计和 Fermi 能级说明载流子浓度.

热平衡 半导体中的自由载流子浓度与导带和价带中允许态的总数有关. 该浓度取决于半导体中的净能量(存在于晶格振动和电子中)。尽管半导体可被外部能量(如入射的光电辐射)所激发,但是在很多情况下总能量只是晶体温度的函数。在这种情况下,半导体会自发地(但不是瞬间地)达到被称为热平衡的状态。热平衡是一种动态平衡,其中的每一过程都与其相反的过程平衡. 例如,热平衡时,如果电子从低能态 E_1 被激发到高能态 E_2,那么同时必然有相同数量的电子从能态 E_2 转移到能态 E_1。同样,如果能量从晶格振动(声子)转移给电子,那么热平衡时会有等量的能量沿相反方向流动。关于热平衡可以理解为一种运动的图像,它可以向前后两个方向运动,但我们看不出图像任何的变化. 下面考虑半导体热平衡时电子和空穴的浓度.

质量作用定律 通常情况下,热能将足以使部分电子从价带激发到导带. 动态平衡时,一部分电子不断地被激发到导带,同时,同样多的电子会失去能量,又跃迁回到价带. 电子从价带到导带的激发过程伴随着一对电子和空穴的产生,而电子通过带隙回到价带时两个载流子同时湮灭,此过程则对应着一对电子-空穴的复合。电子-空穴对的产生率 G 取决于温度 T,在一级近似下与载流子数目无关,可以写出

$$G = f_1(T) \tag{1.1.8}$$

式中 $f_1(T)$ 是由晶体的物理性质和温度决定的函数。另一方面,复合率 R 取决于导带中的电子浓度 n 和价带中的空穴浓度 p(空态),这是因为这两种粒子相互作用时,复合才能发生。因此可将复合率表示为载流子的浓度与包含其他因素的函数 $f_2(T)$ 的乘积

$$R = npf_2(T) \tag{1.1.9}$$

平衡时,产生率必须等于复合率。由等式(1.1.8)和(1.1.9)得

$$npf_2(T) = f_1(T)$$

和

$$np = \frac{f_1(T)}{f_2(T)} = f_3(T) \tag{1.1.10}$$

对给定的半导体材料,等式(1.1.10)表明了热平衡时电子和空穴浓度的乘积只是温度的函数这一重要结论。

在本征(即未掺杂)半导体中,所有载流子都是由越过禁带的激发产生的,因此 $n = p = n_i$,下标 i 表示本征材料。对本征材料,等式(1.1.10)有

$$n_i p_i = n_i^2 = f_3(T) \tag{1.1.11}$$

由于载流子是被热能激发越过禁带而产生的,所以本征载流子浓度取决于温度,同时也是禁带宽度的函数(禁带越宽,能够越过禁带的电子就越少)。下面很快就会看到多数情况下 n_i^2 是由以下表达式给出的

$$n_i^2 = N_c N_v \exp\left(\frac{-E_g}{kT}\right) \tag{1.1.12}$$

式中，N_c和N_v分别是导带边和价带边的等效态密度。尽管N_c和N_v随温度的变化并不明显，但由于温度T还出现在等式(1.1.12)的指数项中，所以n_i对温度的依赖关系是十分明显的。对于硅，$E_g = 1.1\text{eV}$，在室温附近，温度每升高8℃，n_i翻一番。对于给定的半导体，在一定温度下，本征载流子浓度n_i是常数，因此等式(1.1.10)中可用n_i^2代替$f_3(T)$，从而得到一个非常有用的关系式

$$np = n_i^2 \tag{1.1.13}$$

这个式子对本征和非本征半导体都适用。它表明，为增加电子浓度在样品中掺入施主，必将导致样品中空穴浓度的减小，因为只有这样才能使它们的乘积np保持常数。这个结果通常称为质量作用定律，在相互作用的化学粒子之间也存在类似的行为，例如，在酸性或碱性溶液中H^+和OH^-离子的浓度值也满足类似的定理。从前面的推导中可以看出，质量作用定理是产生和复合相平衡，即热平衡的直接结果。

在半导体的中性区（即不存在电场梯度的区域），正电荷的数目必须恰好等于负电荷的数目。正电荷由电离施主和空穴组成，负电荷与电离受主和电子有关[①]。如果假定电中性区所有的掺杂原子都离化

$$N_d + p = N_a + n \tag{1.1.14}$$

利用质量作用定律（等式(1.1.13)）重新写出等式(1.1.14)，得到

$$n - \frac{n_i^2}{n} = N_d - N_a \tag{1.1.15}$$

由此可解出电子浓度n为

$$n = \frac{N_d - N_a}{2} + \left[\left(\frac{N_d - N_a}{2}\right)^2 + n_i^2\right]^{1/2} \tag{1.1.16}$$

在n型半导体中$N_d > N_a$。从等式(1.1.16)中可以看出电子浓度取决于施主与受主浓度之差。这样，正如在前面的例题中看到的，在一块含有受主浓度为N_a的p型材料中，如果掺入更多的施主使得$N_d > N_a$，则变为n型材料。在第2章中将看到制备硅集成电路时这种转变是如何进行的。

硅在室温下的n_i等于$1.45 \times 10^{10}\text{cm}^{-3}$，n型硅中典型的净施主浓度为$10^{15}\text{cm}^{-3}$或更大，因此$(N_d - N_a) \gg n_i$，可将等式(1.1.16)简化为$n \approx (N_d - N_a)$。因此，根据等式(1.1.13)得

$$p = \frac{n_i^2}{n} \approx \frac{n_i^2}{N_d - N_a} \tag{1.1.17}$$

这样，当$N_d - N_a = 10^{15}\text{cm}^{-3}$，有$p = 2 \times 10^5\text{cm}^{-3}$，少数载流子浓度比多数载流子浓度几乎低10个数量级。在非本征半导体中，通常一种载流子浓度比另一种载流子浓度会高出很多个数量级。

Fermi能级　任何宏观半导体材料中，自由载流子（电子或空穴）的数目是相当大的，因而

---

① 在后面的章节中，除非特别指出，讨论器件性能时提到的电子都是指导带中的电子，空穴是指价带中的空状态。

通常可以用统计力学的定律来描述其物理性质[8]。热平衡时电子在允许态中的分布规则是晶体中电子状态的重要特征。根据 Pauli 不相容原理，电子在允许态上的分布遵守 Fermi-Dirac 统计，用 $f_D(E)$ 表示，具有如下形式

$$f_D(E) = \frac{1}{1 + \exp[(E - E_f)/kT]} \tag{1.1.18}$$

式中，E_f 称为 Fermi 能量或 Fermi 能级，它是一个参考能级。从等式(1.1.18)中可以看出 $f_D(E_f)$ 总是等于 1/2。Fermi-Dirac 分布函数通常简称为 Fermi 函数，描述的是能量为 E 的能级被电子占据的几率。如图 1.10(a)所示，当能量 E 低于 E_f 时，Fermi 函数接近于 1，表明大部分低能态是被电子占据的。高能区的 Fermi 函数很小，表明热平衡时几乎没有电子出现在高能级上，这与主观上的直觉是一致的。绝对零度时，低于 E_f 的所有能级被填满，而高于 E_f 的所有能级都是空的。在有限温度下，Fermi 函数并不会发生突变；高于 Fermi 能级的一些能态将被电子占据，而 Fermi 能级以下也会出现一些空态。

Fermi 函数只表示能级被占据的几率，但并不说明该能级实际上是否占有电子，因此仅凭它还不能确定给定能级上的电子数。利用量子物理对系统进行分析，可得到态密度与能量的关系，我们把这个关系用 $g(E)$ 表示。本征半导体材料的 $g(E)$ 如图 1.10(b)所示。在禁带中 $g(E)$ 为零($E_c > E > E_v$)，但是 $g(E)$ 在价带($E < E_v$)和导带($E > E_c$)中增加很快。实际上，电子随能量的分布可以这样求出：将小的能量间隔 dE 内对应的态密度 $g(E)$ 与相应的占据几率 $f_D(E)$ 相乘，然后将这个乘积在整个导带中积分就可得到导带中的电子浓度

$$n = \int_{cb} f_D(E) g(E)\, dE \tag{1.1.19}$$

类似地，价带中空穴的浓度可通过价带的态密度函数与这些态为空态时的几率$[1 - f_D(E)]$的乘积在整个价带积分来求出。

在掺杂浓度不太高的 n 型材料中，导带中只有一小部分允许态被占据，这时 Fermi 函数非常小，且 Fermi 能级远低于导带边。由于$(E_c - E_f) \gg kT$，等式(1.1.18)表示的 Fermi 函数可简化为数学上更简单的 Maxwell-Boltzmann 分布

$$f_M(E) = \exp\left[\frac{-(E - E_f)}{kT}\right] \tag{1.1.20}$$

如果忽略 Pauli 不相容原理对电子在各能级上排列的限制，也可独立地推出这个热平衡函数；即 Boltzmann 函数适用于某一允许能级上可容纳任意数量电子的情况。当能级远高于 Fermi 能级时，它被电子占据的可能性将非常少，以至于 Pauli 不相容原理的限制没有实际的意义，此时 Maxwell-Boltzmann 统计规律适用。

在积分表达式(1.1.19)中代入等式(1.1.20)，经过几步近似可得到用 Fermi 能级表示的导带中的载流子浓度

$$n = N_c \exp\left[-\frac{(E_c - E_f)}{kT}\right] \tag{1.1.21}$$

---

[8] 然而在某些亚微米尺度的器件中，有源区的掺杂原子很少，其数目的统计起伏可能影响器件的特性。

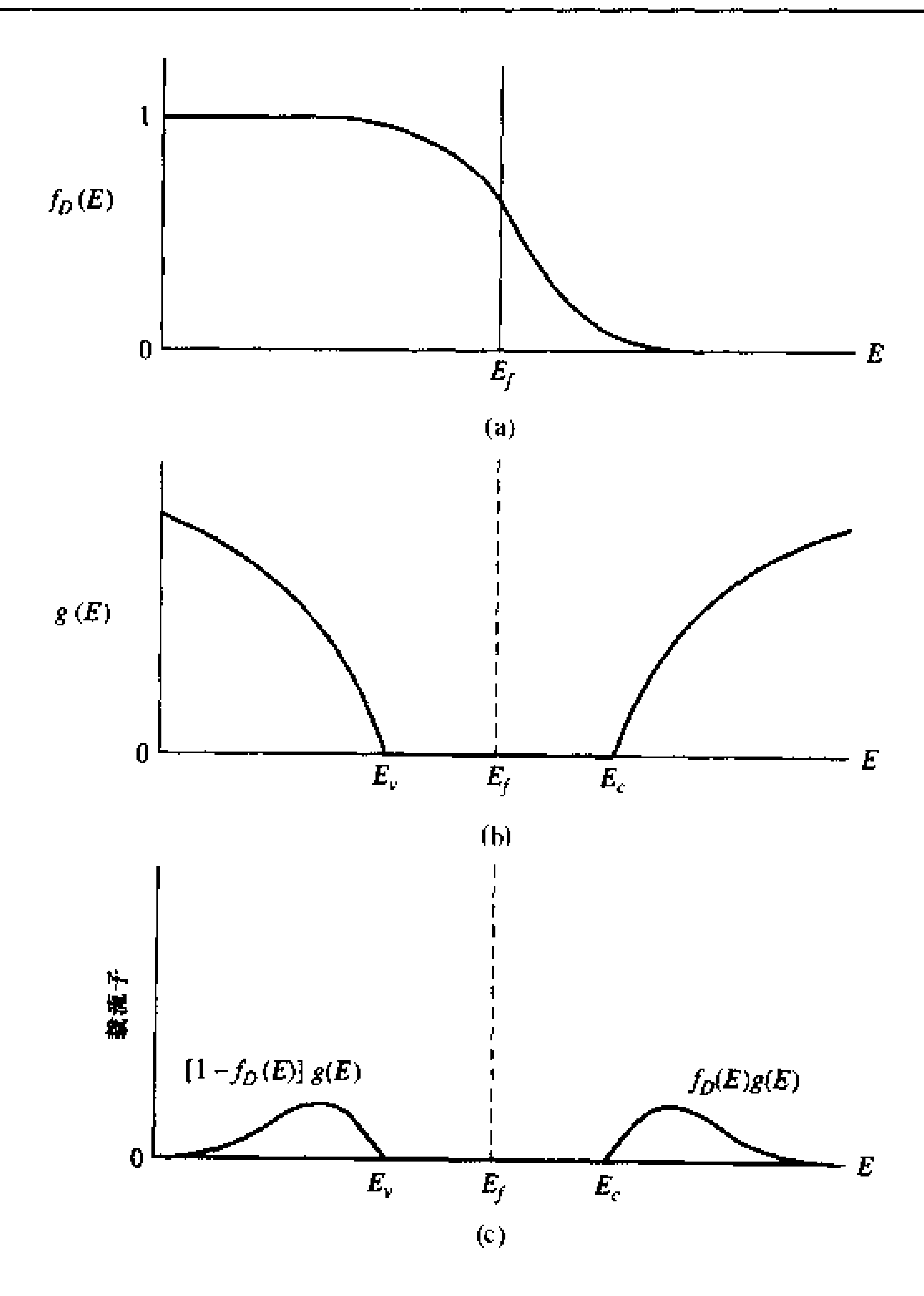


图1.10　(a)Fermi-Dirac分布函数描述能量为E的允许态被一个电子占据的几率;(b)半导体态密度与能量的关系,注意在E_v和E_c之间的禁带中$g(E)$为零;(c)分布函数与态密度的乘积

同样,在中等掺杂的p型材料中,Fermi能级远高于价带边

$$p = N_v \exp\left[-\frac{(E_f - E_v)}{kT}\right] \tag{1.1.22}$$

式中,$(E_c - E_f)$是导带边与Fermi能级之间的能量差,$(E_f - E_v)$是Fermi能级与价带边之间的能量差。N_c和N_v分别称做导带边和价带边的等效态密度,其表达式分别为

$$N_c = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \tag{1.1.23}$$

和

$$N_v = 2\left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2} \tag{1.1.24}$$

式中 m_n^* 和 m_p^* 为电子和空穴的有效质量。这些有效质量与等式(1.1.7)中引入的 m^* 有关，但因为能带的精细结构不同而不完全相同。正如在等式(1.1.21)和(1.1.22)中看到的，所有导带或价带能级的作用可看做是集中在带边 E_c 和 E_v 处，等效态密度就是 N_c 和 N_v。无论 Fermi 能级与带边的间距是几个 kT 或更大，都可以用这些公式来计算热平衡时的载流子浓度。

在等式(1.1.23)和(1.1.24)中除了有效质量 m^* 略微不同外，其他参数都相同，所以有 $N_c \simeq N_v$。因此在 $n \gg p$ 的 n 型材料中，$(E_c - E_f) \ll (E_f - E_v)$，这就意味着 Fermi 能级更接近于导带而远离价带。同样在 p 型半导体中，Fermi 能级更接近于价带而远离导带。

在本征半导体中，$n = p$。因此 $(E_c - E_f) \approx (E_f - E_v)$，此时 Fermi 能级几乎位于禁带中央，$[E_f = (E_c + E_v)/2]$，通常用 E_i 表示本征 Fermi 能级。我们知道 n_i 在处理非本征半导体中载流子浓度(等式(1.1.13))时也很有用，因而 E_i 也常可被用做参考能级

$$n_i = N_c \exp\left[\frac{-(E_c - E_i)}{kT}\right] = N_v \exp\left[\frac{-(E_i - E_v)}{kT}\right] \tag{1.1.25}$$

这时，非本征半导体中的载流子浓度 n 和 p 的表达式(等式(1.1.21)和等式(1.1.22))，可用本征载流子浓度和本征 Fermi 能级重写为

$$n = n_i \exp\left[\frac{(E_f - E_i)}{kT}\right] \tag{1.1.26}$$

和

$$p = n_i \exp\left[\frac{(E_i - E_f)}{kT}\right] \tag{1.1.27}$$

这样，从 Fermi 能级到本征 Fermi 能级间的能量差便是半导体材料掺杂程度的一个度量。因为在 n 型半导体中 E_f 高于 E_i，所以有 $n > n_i > p$，与前面得到的结果一致。

当半导体中含有大量的杂质 $[N_d \to N_c$ 或 $N_a \to N_v$，(对硅 $\sim 10^{19}\,\text{cm}^{-3}$)$]$ 时，就不能再忽略 Pauli 不相容原理的作用了，此时 Fermi-Dirac 分布不能被近似为 Maxwell-Boltzmann 分布。等式(1.1.21 ~ 22)和(1.1.26 ~ 27)不再成立，必须采用更精确的表达式或者对近似表达式的使用范围进行限制。掺杂浓度非常高的半导体($N_d \gtrsim N_c$ 或 $N_a \gtrsim N_v$)称为简并半导体，因为此时 Fermi能级位于导带或价带内。因此电子的允许态与 Fermi 能级的间距非常小，与金属的情况一样。因此，高掺杂半导体材料的许多电学特性与金属相近。

例题　热平衡统计

300K 时，如果在硅材料中含有 $8 \times 10^{16}\,\text{cm}^{-3}$ 的 As 原子和 $2 \times 10^{16}\,\text{cm}^{-3}$ 的 B 原子，计算热平衡时电子和空穴的浓度及 Fermi 能级的位置(相对于本征 Fermi 能级 E_i)。

解：因为施主(As)浓度大于受主(B)浓度，晶体表现为 n 型。净掺杂浓度为施主浓度 8×10^{16} 和受主浓度 2×10^{16} 之差，等于 $6 \times 10^{16}\,\text{cm}^{-3}$。

电子浓度等于净的施主浓度

$$n = 6 \times 10^{16}\,\text{cm}^{-3}$$

空穴浓度(根据等式(1.1.13))为

$$p = \frac{n_i^2}{n} = 3.5 \times 10^3\,\text{cm}^{-3}$$

根据等式(1.1.26)

$$
\begin{aligned}
E_f - E_i &= kT\ln(n/n_i) \\
&= 0.0258\ln(6\times10^{16}/1.45\times10^{10}) \\
&= 0.393\ \text{eV}
\end{aligned}
$$

注意 Fermi 能级相对于导带的值也可用等式(1.1.21)求出

$$
\begin{aligned}
E_c - E_f &= kT\ln(N_c/n) \\
&= 0.0258\ln(2.8\times10^{19}/6\times10^{16}) \\
&= 0.159\ \text{eV}
\end{aligned}
$$

这两个能量之和等于0.55eV，是硅禁带宽度的一半。

↓0.159　E_c
　　　　E_f
↕0.393　E_i

非均匀掺杂半导体　热平衡时，电子遵守 Fermi-Dirac 分布，在给定温度下分布函数由 Fermi 能量决定(等式(1.1.18))。此外，Fermi 能量在整个系统中必须具有统一的值，以保证热平衡时电子传输要求达到的细致平衡。这一非常重要的条件在3.1节和4.1节中还要进行更详细的讨论，这里只考虑整个系统 Fermi 能级一致时对半导体能带的影响。

在例题中看到，未掺杂(本征)半导体中的 Fermi 能级位于禁带中央，非本征(掺杂)半导体的 Fermi 能级更靠近导带或价带[掺施主杂质(n 型材料)时接近于导带，掺受主杂质(p 型材料)时接近于价带]。

画出非均匀掺杂半导体材料的能带图可以说明一些有用的概念。例如图1.11(a)，硅材料中 $x<a$ 的区域内施主浓度为 N_{d1}，在 $x=a$ 时施主浓度锐减为 N_{d2}。当两个区域的半导体密切接触时，整个半导体成一个单一的系统。在绘制这个系统的能带图时，首先要注意 Fermi 能级在热平衡时必须一致，然后根据其他约束条件在 Fermi 能级附近划出导带边和价带边。根据硅的掺杂浓度可确定出远离界面 $x=a$ 处的电子浓度。在 $x\ll a$ 的区域，$n\approx N_{d1}$，在 $x\gg a$ 的区域，$n\approx N_{d2}$。在靠近界面 $x=a$ 的很小的区域内，电子浓度在两者之间变化。

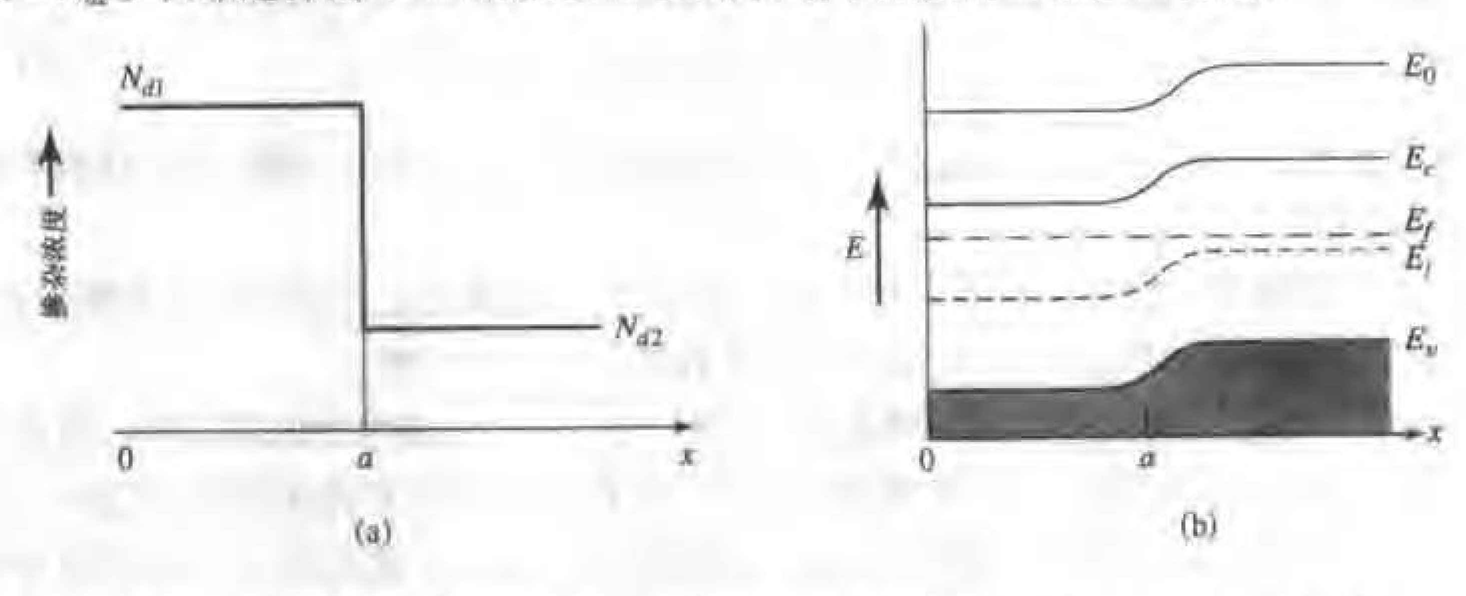


图1.11　(a)硅中的掺杂浓度分布；(b)热平衡时的能带图，已知 n 型半导体中，$0<x<a$ 时施主浓度为 N_{d1}，$x>a$ 时施主浓度为 N_{d2}，其中 $N_{d1}>N_{d2}$

这种非均匀掺杂系统的能带图如图1.11(b)所示。如前所述，我们先画出一水平线表示

统一的 Fermi 能级。在远离界面处,根据电子浓度等于施主浓度,可利用等式(1.1.26)求出每一区域 Fermi 能级和导带边之间的能量差 $E_c - E_f$,从而确定每个区域的导带边。然后通过 $x = a$的界面光滑地画出两个区域间能带的过渡。在过渡区带边的精确变化关系将在第4章中讨论。因为整个半导体材料中 E_g是常数,所以画出的导带边与价带边平行。

从图1.11(b)中可看出,硅中的导带边和价带边沿 x 方向不为常数,施主浓度下降时能带向高能方向移动。导带和价带之间的能量差(禁带宽度)是硅晶体的一个特征参量,不会因晶体轻掺杂或中度掺杂而发生变化(重掺杂的情况在这一节后半部分讨论)。硅带边能量增加意味着掺杂较轻区域的电子势能增加[9]。关于图1.11(b)中的这个问题及其他相关问题将在第4章中考虑。这里的重点是根据 Fermi 能级的一致性,以及利用热平衡原理构造如图1.11(b)所示的能带图。

准 Fermi 能级[†] 前面已经知道 Fermi 能级在解释半导体材料特性时是一个非常有用的概念,下面的讨论推广到器件,我们将会看到 Fermi 能级有更多的应用。Fermi 能级的概念来自热平衡时电子的统计分布,事实上,只有热平衡时 Fermi 能量才有基本的物理定义。然而,辐射引起的激发或者 pn 结上的偏压会破坏热平衡状态。为了分析这些非平衡状态,这里引入两个称做准 Fermi 能级的参数[10]。

我们按照热平衡时电子和空穴浓度与本征载流子浓度之间的关系式(1.1.26)和(1.1.27)来定义准 Fermi 能级。在非平衡条件下,只有定义了电子和空穴不同的准 Fermi 能级后才能使用类似的公式。

定义电子的准 Fermi 能级为 E_{fn}(对应的准 Fermi 势 $\phi_{fn} = -E_{fn}/q$),空穴的准 Fermi 能级为 E_{fp}(对应的准 Fermi 势 $\phi_{fp} = -E_{fp}/q$),且

$$E_{fn} = E_i + kT\ln(n/n_i) \quad 和 \quad \phi_{fn} = \phi_{fi} - \frac{kT}{q}\ln(n/n_i) \tag{1.1.28}$$

和

$$E_{fp} = E_i - kT\ln(p/n_i) \quad 和 \quad \phi_{fp} = \phi_{fi} + \frac{kT}{q}\ln(p/n_i) \tag{1.1.29}$$

式中,ϕ_{fi}是与 E_i相关的 Fermi 势,$\phi_{fi} = -E_i/q$。在非平衡条件下,np 的乘积不等于热平衡值 n_i^2,而是两个准 Fermi 能级之差的函数。从等式(1.1.28)和(1.1.29)可推出

$$np = n_i^2 \exp[(E_{fn} - E_{fp})/kT] \tag{1.1.30}$$

因此,两个准 Fermi 能级之差是半导体中自由载流子偏离热平衡的一个度量,在热平衡条件下这个差值为零。

准 Fermi 能级的概念在分析光电导时特别有用,这时过剩电子和空穴是由光激发产生的。在第5章中将会看到,用准 Fermi 能级讨论产生和复合也非常方便。

光电导[†] 如果入射光子能量足够高,就可能打破晶格中束缚电子的共价键。当这个键断裂时,无论自由电子还是留下的空键都可以作为电流的载体在半导体中运动。这种自由载流子的产生,称做光致激发;如果用能带图描述,就是将电子从价带激发到导带,在价带中留下空

---

⑨ 提供了电子的施主原子形成一个局域的正电荷。高的施主浓度(因而高的正离子浓度)趋向于吸引电子到这一区域来,因此降低了该区域的电子势能。

⑩ 有些作者用 Imref,即将 Fermi 反过来拼写表示准费米能级,意思是虚拟参考能级。

穴。光致激发所需的光子能量至少要等于禁带宽度，且产生的空穴数等于电子数。硅的带隙(1.1eV)对应于电磁波谱中远红外光子的能量(波长1.1μm)。

照射到半导体表面的光进入晶体内部时会被吸收。沿入射方向小的长度增量 Δx 上被吸收的能量 ΔI 可用吸收系数 α 表示

$$\Delta I = I(x) - I(x + \Delta x) = I(x) \times \alpha \Delta x \tag{1.1.31}$$

式中 $I(x)$ 是入射光到达 x 点时的能量。取 Δx 为无穷小量时，等式1.1.31可重新写成微分方程的形式，它的解为

$$I(x) = I_0 \exp(-\alpha x) \tag{1.1.32}$$

式中 I_0 是 $x=0$ 处进入固体的入射光能量。

吸收系数与光子能量有较强的依赖关系，图1.12给出了硅的吸收系数与波长(光子能量)[①]的关系曲线。高能紫外(UV)光在硅晶体中的特征长度(等于 α^{-1})小于10nm，而波长为1μm(自由空间)的光在硅中则不会被有效地吸收，其穿透深度可达100μm。大于禁带能量的光子被吸收后几乎全部用来产生电子和空穴。光吸收曲线的特殊形状与硅的能带结构有关，关于这个重要问题在固体物理中有更详细的讨论。

当硅材料中发生光致激发时，入射光能转化为晶格热能。此时，硅处于非平衡态，所以用准Fermi能级来描述自由载流子浓度是恰当的。

例题 光激发和准 Fermi 能级

已知硅片中施主掺杂浓度为 10^{15}cm^{-3}。

(a)求电子和空穴的浓度及Fermi能级相对于本征Fermi能级的位置。

(b)光照射硅片后，产生的稳态光生电子和空穴的浓度为 10^{12}cm^{-3}。假设硅片厚度小于光的吸收长度，因此自由载流子在整个体积中均匀地产生。求硅片中总的电子和空穴浓度并计算两种载流子的准Fermi能级的位置。

(c)将光强增加，使得光激发产生的电子-空穴对浓度达到 10^{18}cm^{-3}，重复计算(b)。

解：

(a) $n = N_d = 10^{15}\text{cm}^{-3}$

$p = \dfrac{n_i^2}{n} = 2.1 \times 10^5 \text{cm}^{-3}$

$E_f - E_i = kT\ln(n/n_i) = 0.29\text{eV}$

(b) $n = 10^{15} + 10^{12} \approx 10^{15}\text{cm}^{-3}$

$p = 2.1 \times 10^5 + 10^{12} \approx 10^{12}\text{cm}^{-3}$

$E_{fn} - E_i = kT\ln(n/n_i) = 0.29\text{eV}$

$E_i - E_{fp} = kT\ln(p/n_i) = 0.11\text{eV}$

(c) $n = 10^{15} + 10^{18} \approx 10^{18}\text{cm}^{-3}$

$p = 2.1 \times 10^5 + 10^{18} \approx 10^{18}\text{cm}^{-3}$

---

① 电磁波波长 λ 与光子能量 E 之间的关系为 $\lambda = hc/E$，其中 hc 为普朗克常数与光速的乘积。如果 λ 的单位为微米，E 的单位为eV，则转换关系为 $\lambda = 1.24/E$。

$$E_{fn} - E_i = kT\ln(n/n_i) = 0.47\text{eV}$$

$$E_i - E_{fp} = kT\ln(p/n_i) = 0.47\text{eV}$$

在(b)问中可以看出，光激发使得少子浓度增加了 7 个数量级，而多子浓度并没有明显的变化。因此，电子的准 Fermi 能级接近于热平衡时的 Fermi 能级，但是空穴的准 Fermi 能级移动了 0.40eV。在(c)问中光的强度增加后，电子和空穴浓度都受到影响，两者的准 Fermi 能级都远离热平衡时的位置。这种情况下两种载流子浓度几乎相等，与本征半导体在高温时的情况类似。掺杂半导体中，光激发的大部分情况与本例 (b)问中的情形相似，即光激发大大改变了少子浓度，而多子浓度基本不受影响。

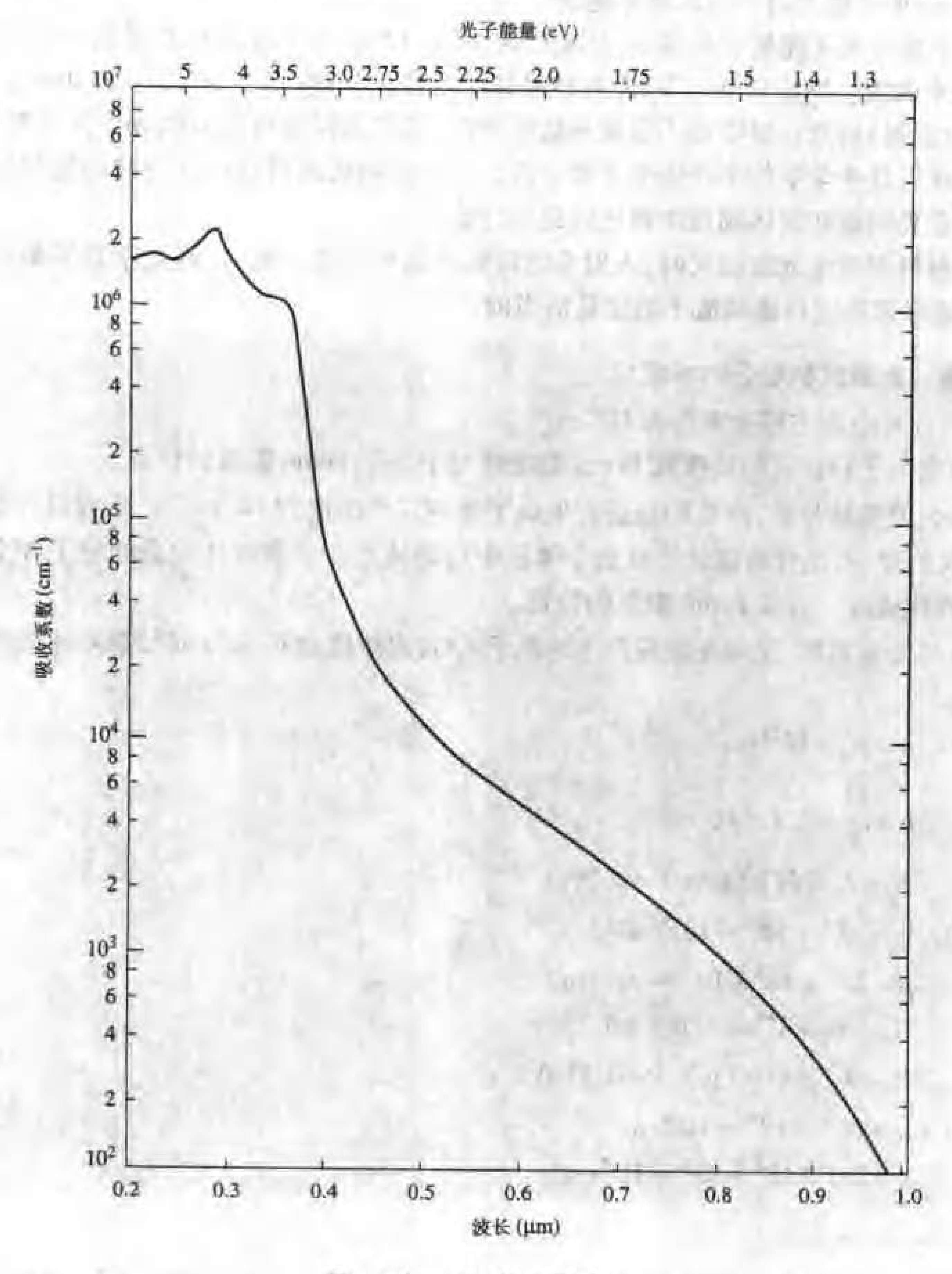


图 1.12 硅的光吸收系数

重掺杂† 前面讨论的大部分情况都假定导带中只有少部分能级被电子占据，价带中只有少量的能级是空的，从而简化了半导体中电子和空穴浓度的统计表达式。在这些假定下，例如，我们可以对等式(1.1.19)的积分式进行近似，定义"导带边等效态密度"N_c，并用更简单的Maxwell-Boltzmann统计近似代替Fermi统计来算电子浓度(等式(1.1.21))。然而在晶体掺杂浓度接近N_c时这些近似不再适用。半导体重掺杂时，必须考虑其他更基本的效应才能正确计算出自由载流子的统计分布。

对中等掺杂(例如硅片中的体掺杂)，杂质原子彼此间没有相互作用，它们并不影响晶体的能带结构。例如，掺杂浓度为$5\times10^{15}\text{cm}^{-3}$时每$10^7$个硅原子中只有一个杂质原子。每个杂质原子在硅禁带中增加一个分立的允许的施主能级。如果杂质浓度增加到与硅原子浓度相比不可忽略时，能带结构本身开始受到影响。

最重要的影响是硅禁带宽度降低，从而引起自由载流子浓度n和p的乘积增加。这一效应通常用pn乘积以下面的形式表示出来

$$pn = n_i^2\exp(\Delta E_g/kT) = n_{ie}^2 \tag{1.1.33}$$

式中，ΔE_g表示由于重掺杂引起的有效禁带宽度的降低，n_{ie}是本征载流子浓度的有效值。测量表明，当掺杂浓度小于10^{18}cm^{-3}时可忽略带隙的变化，但是在更高的掺杂浓度下，带隙的变化会相当大。图1.13给出了硅中ΔE_g随自由电子浓度n变化的一些实验数据。重掺杂效应在掺杂浓度达到$10^{18}\sim10^{19}\text{cm}^{-3}$时比较明显；当电子浓度达到$10^{19}\text{cm}^{-3}$时，$\Delta E_g$超过带隙能量的10%。

详细的研究表明，随着掺杂浓度的增加，引入的杂质能级将不再是分立的，而是展宽为能带。这些杂质能带可与邻近的导带或价带交叠，这时，杂质原子不需要能量就可以离化产生自由载流子。因此，在重掺杂情况下，本章前面推出的有关公式需要修正。

重掺杂对器件最重要的影响是，限制了双极晶体管的电流增益，还会在双极和MOS器件中增加不希望有的漏电流。

1.2 半导体中的自由载流子

本章开始时首先提到了固体的电学特性，当样品两端加电压时，流过的电流与电压之间通常表现为大家熟悉的线性关系。这一关系被称为欧姆定律：$V=IR$。尽管欧姆电导的详细的物理过程的推导可能会非常复杂，但是对这一过程的近似描述就能满足我们的需要。为此，我们先介绍不加外电场时自由电子运动的图像，然后考虑在低、中等电场强度下许多器件应用中的特征，最后考虑高电场强度下的情况。

首先回顾一下半导体中的电子(和空穴)，它们几乎是"自由粒子"，不属于任一特定的晶格位置。晶格势场作用力对电子和空穴的影响体现在有效质量中，使它们的有效质量与自由电子质量略有不同。根据统计力学规律，可以写出电子和空穴具有的与经典自由粒子相应的热能：每个自由度具有$kT/2$的能量，其中k为Boltzmann常数，T为绝对温度。这意味着晶体中的电子在有限的温度下不是静止的，而是以随机的速度运动着。此外，电子的均方热运动速

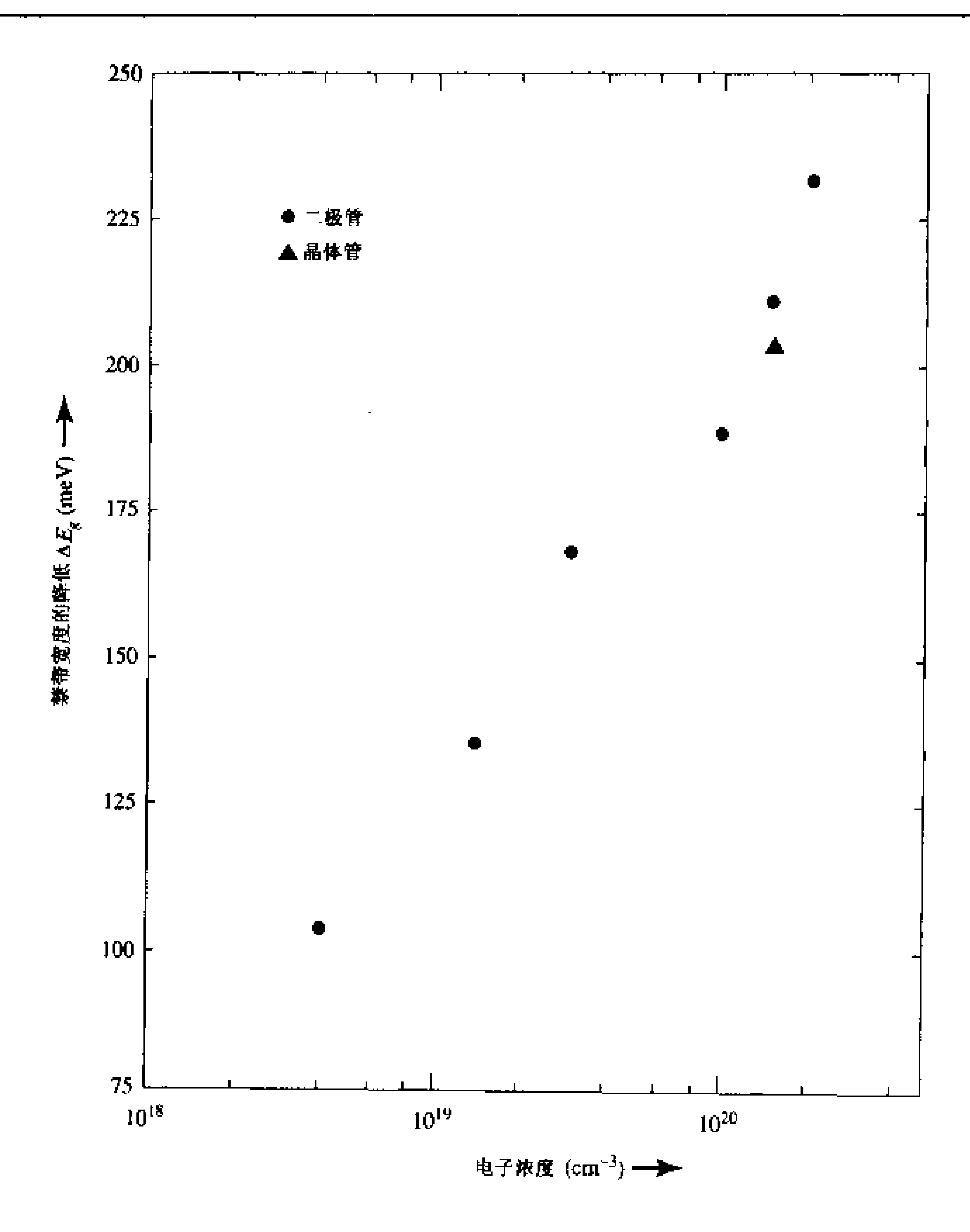


图 1.13 禁带宽度的降低 ΔE_g 与电子浓度的关系。[A. Neugroschel, S. C. Pao 和 F. A. Lindholm, IEEE Trans. Electr. Devices, ED-29, 894(May 1982)]

度v_{th}与温度之间的近似[①]关系为

$$\frac{1}{2}m_n^* v_{th}^2 = \frac{3}{2}kT \tag{1.2.1}$$

式中 m_n^* 为导带电子的有效质量。对晶体硅，$m_n^* = 0.26m_0$（m_0为自由电子的静止质量），$T = 300K$ 时从等式(1.2.1)中计算出的 $v_{th} = 2.3 \times 10^7 cm\ s^{-1}$。可以认为电子在晶格中沿随机方向运动，电子与电子、电子与晶格之间会发生相互碰撞。热平衡时，电子的运动完全是随机的，所以沿任何方向的净电流均为零。电子与晶格之间的碰撞，导致电子和构成晶格的原子核之间

① 由于取平均的方式不太恰当，等式(1.2.1)略有误差。但是，我们只关心结果的数量级。300K 时，硅中的电子或空穴的 v_{th} 通常取 $10^7 cm\ s^{-1}$。

的能量交换。电子的平均散射时间用τ_{cn}表示，它是将每两次碰撞之间的时间间隔对所有的电子取平均后得到的。以上讨论适用于没有外加电场，热平衡态的晶体。

1.2.1 漂移速度

现在对晶体施加一小电场。电子在两次碰撞之间的时间间隔内沿着电场力的方向被加速。图1.14(a)是典型的外加电场$\mathscr{E}$较小时晶体中电子的运动图。注意图中电子沿电场力方向的运动只是叠加在随机热运动速度上的一个小的扰动。因此，平均散射时间τ_{cn}并不因外加电场的影响而发生明显变化。

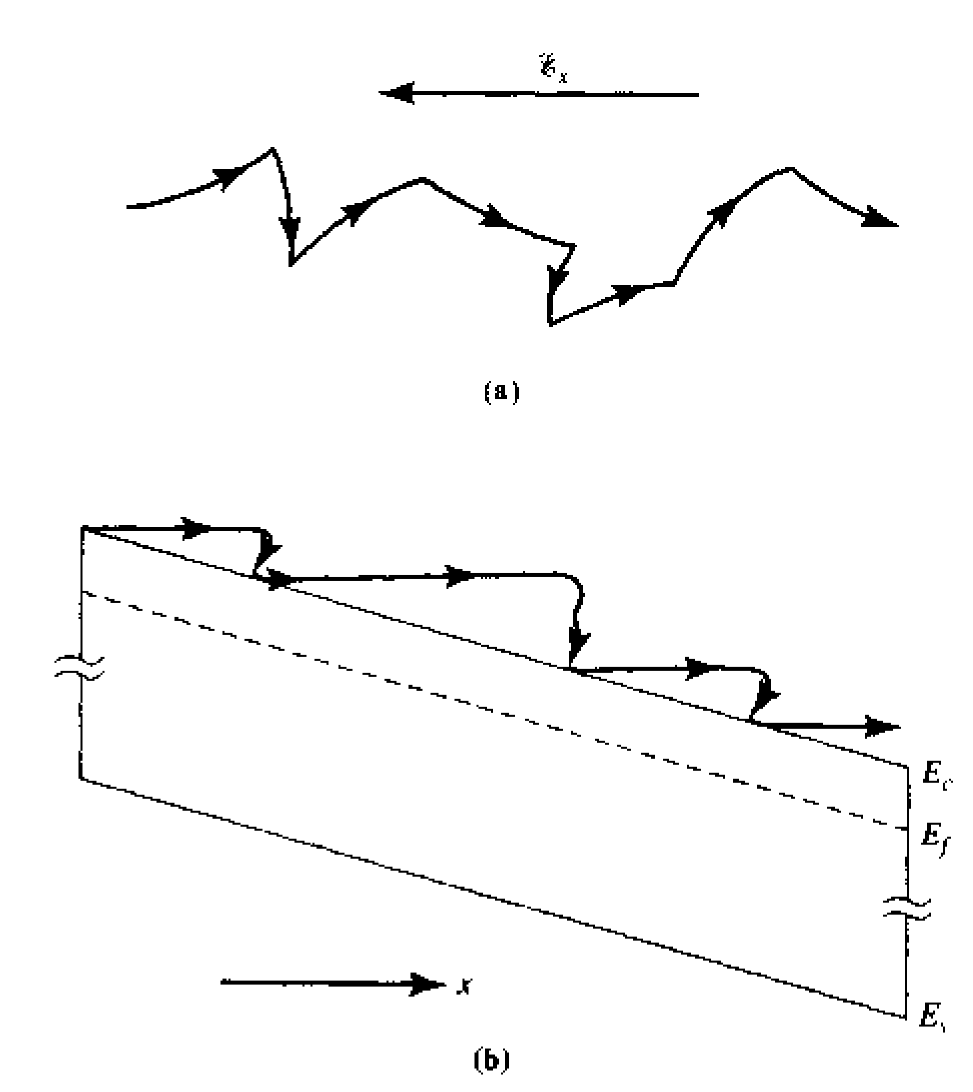


图1.14 (a)外加电场对固体中电子运动的影响；(b)电子运动的能带示意图，表示电子在经历一次碰撞后的能量损失

图1.14(b)给出了小的外加电场下电子运动的能带示意图。当外加电场为常数时，晶体中的能级线性变化。如果外加电场沿x负方向，电子(能带图向下倾斜)趋于向图的右端运动。当电子与晶格碰撞时交换能量，电子的能量下降至接近于热平衡时的值。如果电场较小，交换的能量也少，晶格没有被流过的电流明显加热。图1.14(b)中给出放大后的能带的斜率以及与晶格碰撞相关的能量损失。事实上，每一次碰撞造成的能量损失远小于电子的平均热能。导带中静止的电子位于带边E_c。因此，导带中电子的动能为$(E-E_c)$，热平衡时的平均值只有$(E-E_c)=\frac{3}{2}kT$，或者说在300K时为0.04eV。这一值不到带隙能量的4%。

外加电场作用下载流子获得的净速度称为漂移速度v_d。根据电子在两次碰撞之间获得的冲量(作用力乘以时间)与同一时间内获得的动量相等，可求出v_d；这是因为稳态时在两次碰撞之间获得的所有动量在碰撞时都交给了晶格。作用在电子上的力为$-q\mathscr{E}$，获得的动量为

$m_n^* v_d$。这样(在等式(1.2.1)脚注中提到的统计范围内)

$$-q\mathscr{E}\tau_{cn} = m_n^* v_d \tag{1.2.2}$$

或

$$v_d = -\frac{q\mathscr{E}\tau_{cn}}{m_n^*} \tag{1.2.3}$$

等式(1.2.3)表明电子漂移速度 v_d 与电场成正比,比例因子与平均散射时间和近自由电子的有效质量有关。这个比例因子是电子的一个重要特征参量,称为迁移率,用符号 μ_n 表示。

$$\mu_n = \frac{q\tau_{cn}}{m_n^*} \tag{1.2.4}$$

因为 $v_d = -\mu_n\mathscr{E}$,迁移率描述的是电子在外加电场作用下运动难易的程度。

根据等式(1.2.3),沿外加电场方向的电流密度可通过对单位体积内每个电子的电荷与速度的乘积求和得到

$$J_n = \sum_{i=1}^{n} - qv_i = -nqv_d = nq\mu_n\mathscr{E} \tag{1.2.5}$$

前面的讨论完全适用于空穴。价带中空穴的动能用$(E_v - E)$表示,位于价带边 E_v 的空穴动能为零。如果带边是倾斜的,空穴在电子能带图中向上运动。迁移率 μ_p 定义为 $\mu_p = q\tau_{cp}/m_p^*$。总电流为电子电流和空穴电流之和

$$J = J_n + J_p = (nq\mu_n + pq\mu_p)\mathscr{E} \tag{1.2.6}$$

等式(1.2.6)中圆括号内的表达式定义为半导体的电导率 σ

$$\sigma = q\mu_n n + q\mu_p p \tag{1.2.7}$$

在非本征半导体中,因为两种载流子浓度差异较大,等式(1.2.7)中只有一项起主要作用。电阻率,即电导率的倒数,与掺杂浓度之间的关系示于图1.15。图中给出的是掺P的n型硅和掺B的p型硅两种情况。对不同的掺杂元素电阻率会略有不同,尤其是在重掺杂区。但是,大部分情况下,图1.15可用于任何掺杂元素。

固体中与电导率密切相关的一个特征值是介电弛豫时间,描述的是半导体中的电荷通过导电过程消散,从而达到电中性所需的时间。金属中的介电弛豫时间较短,绝缘体和半导体中的较长。根据同一种材料中电荷的渡越时间和介电弛豫时间的相对大小,可很容易解释器件的一些概念。习题1.12对介电弛豫时间给出了进一步的介绍。

1.2.2 迁移率和散射

量子力学计算表明完美的周期性晶格对自由载流子没有散射作用,即载流子不与静止的、完美的晶格交换能量。然而,高于绝对零度时,构成晶格的原子会在平衡位置附近振动。这些振动破坏了晶格周期性,使得载流子与晶格之间有能量交换[13]。载流子与晶格振动之间的相互作用可看做与具有一定能量的被称为声子的"粒子"的碰撞。声子与光子一样,具有以 $h\nu$

[13] 由于这一能量是由外电场提供给载流子的,散射过程会导致半导体发热。热耗散通常是限制半导体器件尺寸的一个因素。一个器件的尺寸必须足够大,以避免温度过高而不能正常工作。

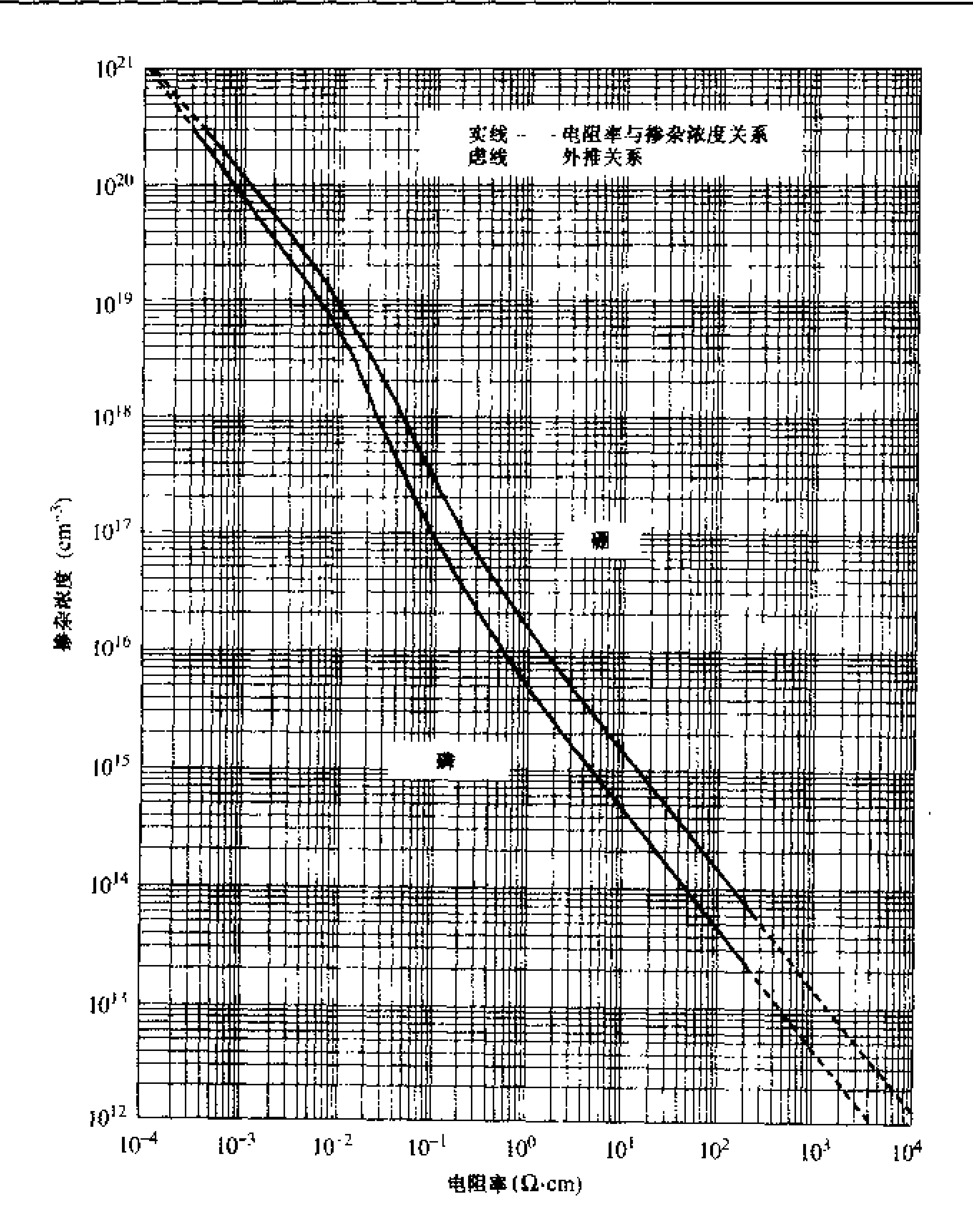


图1.15　23℃(296K)时，掺P和掺B的硅样品中杂质浓度与电阻率的关系曲线，该曲线也可表示300K时的值，只是略有误差[W. R. Thurber, R. L. Mattis, Y. M. Liu, National Bureau of Standfords Special Publication 400 - 64, 42(May 1981)]

为单位的量子化能量，其中ν为晶格振动频率，h为普朗克常数。利用声子的概念，电导率和热导率公式通常可被简化。硅在室温和中等电场强度下，晶格最低的振动模对应的声子能量为0.063eV，当电子与这些最低能量的声子相互作用时，交换的能量就在这个数量级。在更高的温度下，晶格振动变得更剧烈，电子与晶格振动间的作用更为重要。理论分析表明，当晶格散射起主导作用时，迁移率应随温度增加按照T^{-n}关系下降，n在1.5到2.5之间。实验得出的n在1.66到3之间，通常为$n=2.5$。

除了晶格振动，杂质原子也会引起晶格的局部畸变，对自由载流子有散射作用。然而与晶格振动散射不同，电离杂质散射在高温时则变得不太重要。因为载流子在高温时运动速度更快，在杂质原子附近停留的时间更短，因而散射较弱。当杂质散射为主要的散射机制时，迁移率随温度升高而升高。背景杂质和晶格缺陷也会引起散射，这些缺陷或者来自于对半导体的

晶体质量没能很好地控制,或与多晶材料的晶粒间界有关。用在很多 MOS 集成电路中的多晶硅薄膜(第 2 章和第 9 章)就是后者的一个例子。在同样的掺杂条件下,多晶材料中的晶粒间界和缺陷使得迁移率远小于单晶材料。

上面讨论的两种或者更多种的散射过程可能同时起重要作用,那么就必须考虑它们对迁移率的共同影响。为此,首先考虑在时间间隔 dt 内被散射的粒子数。一个载流子在 dt 内被第 i 种机制散射的几率为 dt/τ_i,其中 τ_i 为 i 机制对应的两次散射之间的平均时间。那么,在时间间隔 dt 内一个载流子的总的散射几率 dt/τ_c 为各种散射几率之和

$$\frac{dt}{\tau_c} = \sum_i \frac{dt}{\tau_i} \tag{1.2.8}$$

由于各种散射机制同时起作用,总的平均散射时间小于其中任一机制的平均散射时间,并且散射时间最短的那一种机制起主要作用。由于迁移率 μ 等于 $q\tau_c/m^*$,可写出

$$\frac{1}{\mu} = \sum_i \frac{1}{\mu_i} \tag{1.2.9}$$

在几种不同的散射机制同时起作用时,载流子迁移率的倒数为对应于各种散射机制的迁移率的倒数之和;总的迁移率小于任何一种散射机制所对应的值。因为迁移率各分量的倒数关系(等式(1.2.9)),总的迁移率将主要由 τ_i 最小的散射机制决定。

室温下,考虑这些因素得到的硅中电子和空穴的迁移率如图 1.16 所示。图中的曲线是对许多不同来源的测量数据的一个最好的拟合。在轻掺杂材料中,电离杂质散射所决定的迁移率将大于晶格散射所决定的迁移率值。因此在杂质浓度小于 $10^{15}cm^{-3}$ 时,电子和空穴的迁移率几乎不随掺杂浓度变化。但是,高掺杂浓度下,电离杂质散射的作用与晶格振动散射的作用相当或更大一些,总的迁移率则随掺杂浓度升高而下降。

如果用杂质补偿的方法改变半导体的导电类型(p→n 或 n→p),可观察到迁移率与总的杂质浓度之间的依赖关系。载流子浓度取决于两种类型的杂质浓度差($N_d - N_a$)(等式(1.1.16)和等式(1.1.17)),而散射则与两种电离杂质浓度之和($N_d + N_a$)有关。这样在相同的净载流子浓度的条件下,补偿半导体的迁移率要远低于未补偿材料。

用于拟合图 1.16 中硅的电子和空穴迁移率数据的公式为[3]

$$\mu = \mu_{min} + \frac{\mu_{max} - \mu_{min}}{1 + (N/N_{ref})^{\alpha}} \tag{1.2.10}$$

式中,N 为硅中总的杂质浓度,四个参数 μ_{max},μ_{min},N_{ref} 和 α 对不同的杂质有不同的值。在图 1.16的图注中给出了硅中最常用杂质的有关参数值。表 1.1 列出了不同杂质浓度对应的迁移率(由等式(1.2.10)计算)的数值。

从图 1.16 中可看出,当总的杂质浓度小于 $10^{19}cm^{-3}$ 时,杂质种类对电子迁移率的影响很小。在高掺杂区($N > 10^{19}cm^{-3}$),掺 P 硅样品中的迁移率比掺 As 样品的迁移率大 10%~20%。当掺杂浓度非常高(大于 $10^{20}cm^{-3}$)时,测量得到的迁移率小于图 1.16 中给出的最小值。

温度关系　由前面的简要讨论已知,不同的散射机制对迁移率的影响与温度的关系是不同的。例如,电离杂质对载流子的散射随温度升高会变得不太重要,因为运动速度较快的载流子与静止的杂质间的相互作用会减弱。然而,高温时晶格振动散射(声子散射)则变得更为重要。因此低温时迁移率随温度的升高而增加(杂质散射为主),而高温时迁移率随温度的升高

而降低(声子散射为主)。如图1.17所示的实验数据，不同的温度依赖关系使得迁移率随温度的变化曲线出现极大值。在迁移率极值处，两种机制的迁移率随温度的变化趋势相抵消，迁移率具有最小的温度灵敏度。

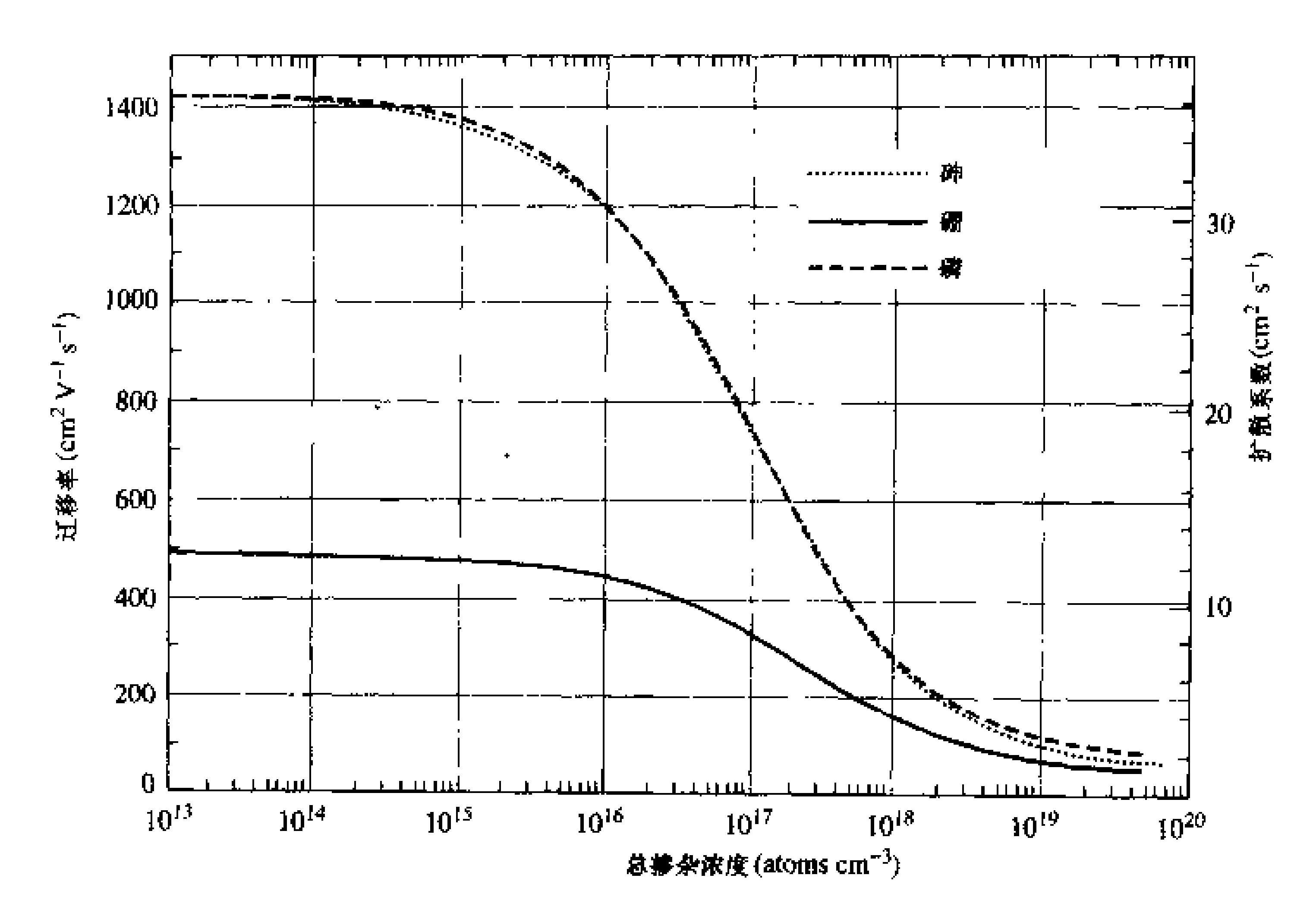


图1.16　300K时硅中电子和空穴的迁移率与总杂质浓度的关系。图中曲线为不同来源的测量数据的拟合结果，迁移率曲线可用以下参数值代入等式(1.2.10)中算出[3]

参数	As	P	B
μ_{min}	52.2	68.5	44.9
μ_{max}	1417	1414	470.5
N_{ref}	9.68×10^{16}	9.20×10^{16}	2.27×10^{17}
α	0.680	0.711	0.719

表1.1　硅中的迁移率($cm^2V^{-1}s^{-1}$)

N	电子		空穴
	砷	磷	硼
10^{13}	1423	1424	486
10^{14}	1413	1416	485
10^{15}	1367	1374	478
10^{16}	1184	1194	444
10^{17}	731	727	328
10^{18}	285	279	157
10^{19}	108	115	72

迁移率与温度和掺杂浓度的关系式可用于器件设计和性能分析。对硅来说，已有如下的经验公式[4]

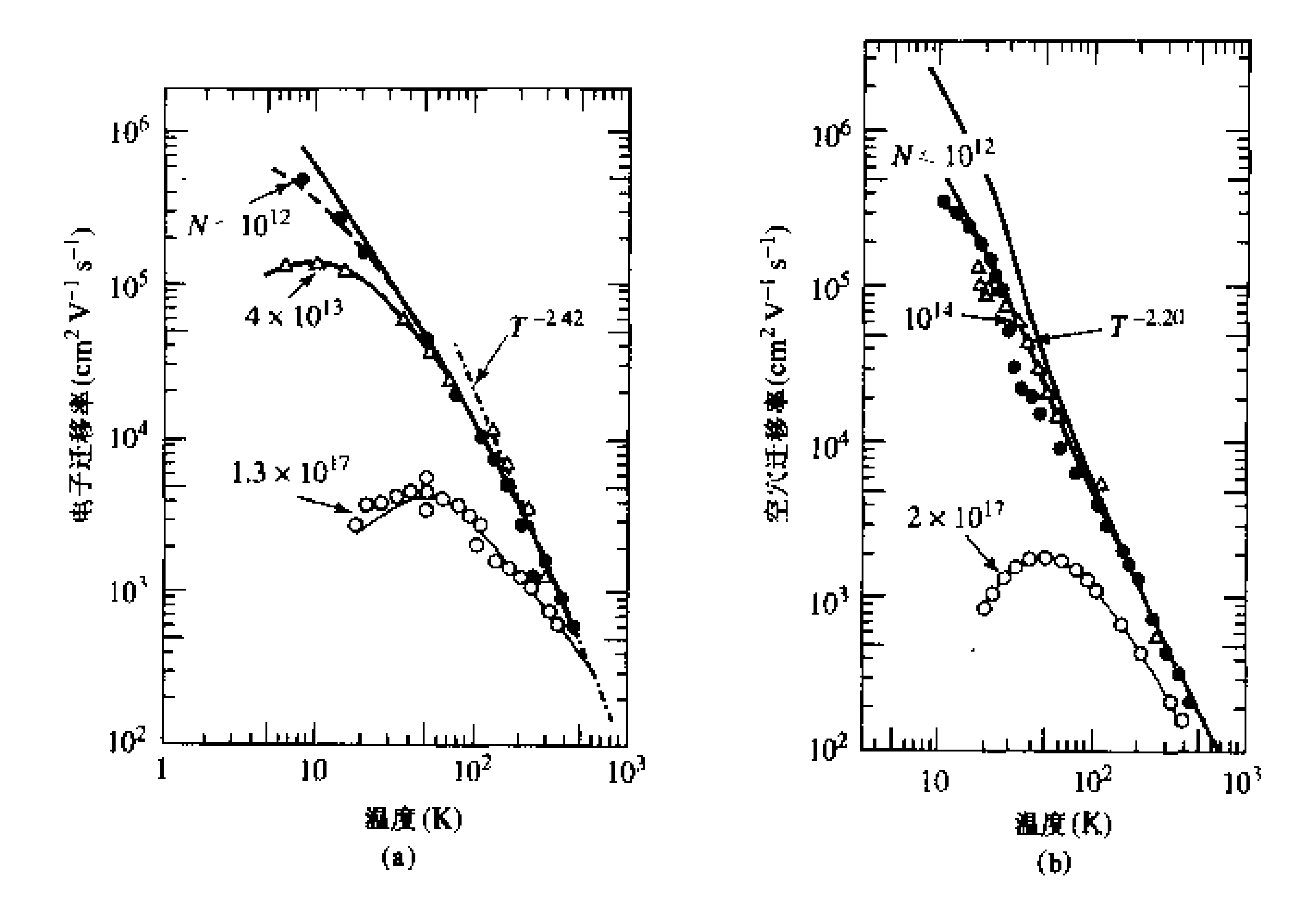


图 1.17 硅中的低电场迁移率与温度的关系。(a)电子;(b)空穴,实线表示晶格散射的理论值[5]

$$\mu_n = 88\, T_n^{-0.57} + \frac{1250\, T_n^{-2.33}}{1 + [N/(1.26 \times 10^{17}\, T_n^{2.4})]0.88\, T_n^{-0.146}}$$

和

$$\mu_p = 54.3\, T_n^{-0.57} + \frac{407\, T_n^{-2.23}}{1 + [N/(2.35 \times 10^{17}\, T_n^{2.4})]0.88\, T_n^{-0.146}} \tag{1.2.11}$$

式中,$T_n = T/300$,T 以 K(Kelvin)为单位;N 是硅中总的杂质浓度。在掺杂浓度小于 10^{20}cm^{-3}、温度为 250~500K 时,等式(1.2.11)是适用的。

速度饱和 到目前为止,我们一直简单地假定(认为 τ_c 对 $\mathscr{E}$ 不敏感)由外加电场赋予载流子的速度远小于随机热运动速度,从等式(1.2.1)中可算出室温下硅中载流子的热运动速度约为 10^7cm s^{-1}。当硅中的电子迁移率取 $\mu_n \approx 1400\ \text{cm}^2\text{V}^{-1}\text{s}^{-1}$,外加场强取典型值 $100\ \text{V cm}^{-1}$ 时,漂移速度约为热运动速度的 1.5%,外加电场并没有显著改变总的载流子的速度或能量。但是,在高电场下,漂移速度将会变得与随机热运动速度相当,从而不能再被看做叠加在热运动上的一个小量。这时载流子的总能量将随外电场的增强而显著增加。当载流子能量超过环境温度对应的热能时,通常称之为热载流子,用有效温度 T_e 表征。随着电场的增强,与载流子动能的增加相对应,T_e 也随之增高[6]。

在强电场下,热电子的能量达到某个临界值时,另一种散射机制(与高能量的光学声子的碰撞)变得很重要,迁移率开始从低电场对应的值下降。因为光学声子散射能够有效地使热载流子将从电场中获得的能量传给晶格,从而载流子能量不再明显升高。这时漂移速度达到强场下的极限漂移速度 v_l 或饱和漂移速度 v_{sat}。图 1.18 给出了测量得到的硅中电子(77K 和 300K)和空穴(300K)的漂移速度与外加电场的关系图。低电压下,曲线表现为线性,表明迁移率为常数。但是,当电场高于每厘米几千伏特时,迁移率明显偏离常数值。因为这一范围的

电场在集成电路器件中很普遍（等价于每微米加几百毫伏的电压），所以在分析很多实际器件时，必须考虑速度饱和的问题。

作为一种有用的近似，图1.18中的数据可用下面的经验公式表示[5]

$$|v_d| = v_l \frac{\mathscr{E}}{\mathscr{E}_c}\left[\frac{1}{1+(\mathscr{E}/\mathscr{E}_c)^{\beta}}\right]^{1/\beta} \tag{1.2.12}$$

等式(1.2.12)中的参数 v_l、$\mathscr{E}_c$ 和 β（为绝对温度 T 的函数）在表1.2中给出。

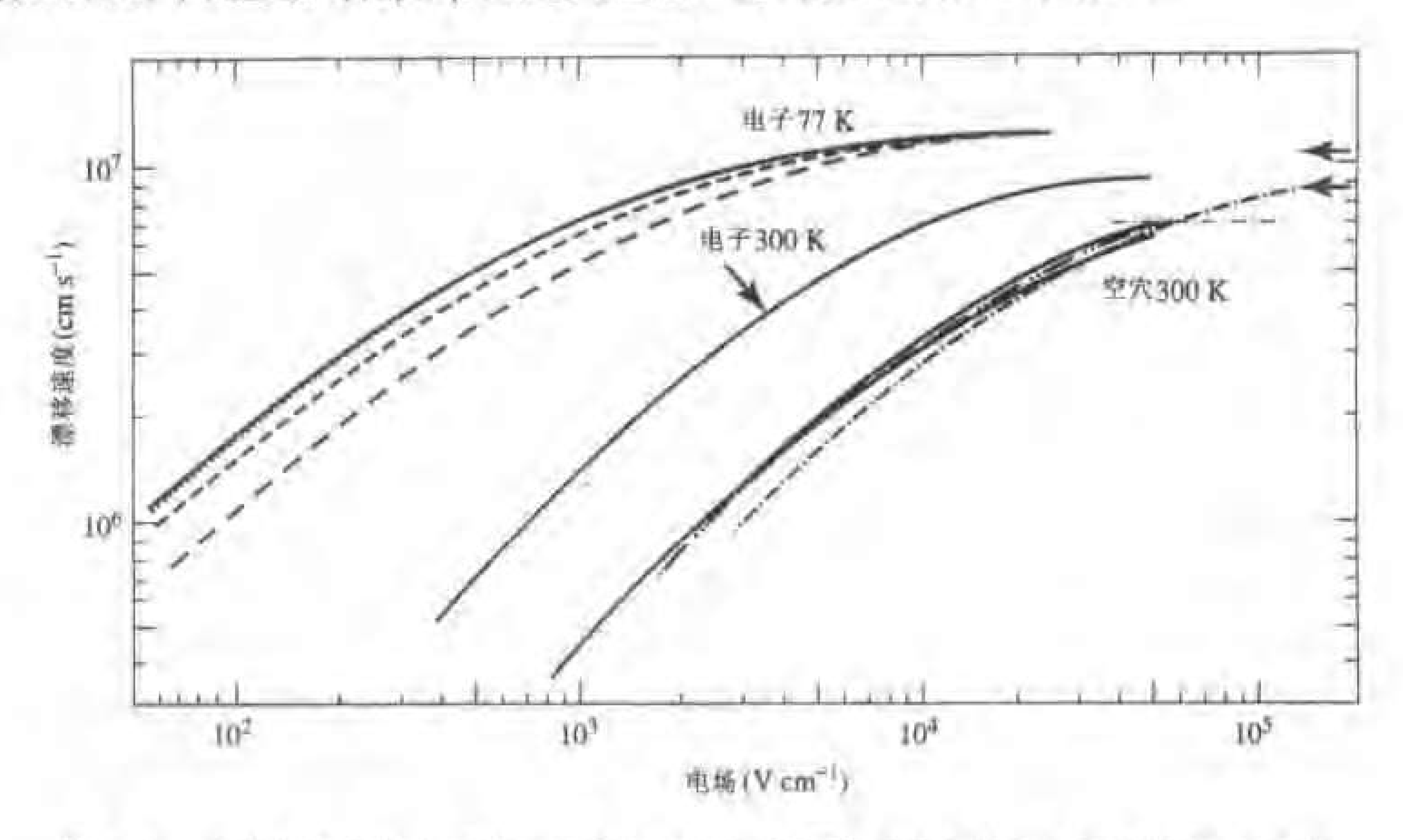


图1.18　硅中电子（77K和300K）和空穴（300K）的漂移速度与外加电场的关系，强电场时速度达到饱和。图中几条曲线对应于报导的不同测量数据，对曲线的最佳拟合公式见等式1.2.12和表1.2[5]

表1.2　计算电场和漂移速度的参数

参　数	电　子		空　穴	
	表达式	300K	表达式	300K
v_l cm s^{-1}	$1.53\times10^9 T^{-0.87}$	1.07×10^7	$1.62\times10^8 T^{-0.52}$	8.34×10^6
$\mathscr{E}_c$ V cm^{-1}	$1.01T^{1.55}$	6.91×10^3	$1.24\times T^{1.68}$	1.45×10^4
β	$2.57\times10^{-2} T^{0.66}$	1.11	$0.46T^{0.17}$	2.637

例题　速度饱和

已知300K时，有效电子迁移率（定义为漂移速度与电场之比）降低到低电场迁移率的一半，利用等式(1.2.12)求此时对应的电场。

解：在低电场下，漂移速度 v_d 与电场成正比（等式(1.2.3)），并且 $|v_d| = \mu_n \mathscr{E}$。低电场条件下($\mathscr{E}/\mathscr{E}_c \ll 1$)，利用等式(1.2.12)可以得到用 $\mathscr{E}_c$ 和 v_l 表示的低场迁移率 $\mu_n = |v_d/\mathscr{E}| = |v_l/\mathscr{E}_c|$。将表1.2中的值代入这个表达式中得到低场迁移率为 $\mu_n = 1548\mathrm{cm^2V^{-1}s^{-1}}$，大约高于图1.16中的值的10%。当参数来自拟合曲线时，缺乏一致性是很正常的。

当有效迁移率减为低场迁移率的一半时,我们注意到 $\mu_n = |v_d/\mathscr{E}| = (1/2) \times |v_l/\mathscr{E}_c|$。因此根据等式(1.2.12)有

$$\frac{1}{2} = \left[\frac{1}{1 + (\mathscr{E}/\mathscr{E}_c)^\beta}\right]^{1/\beta}$$

式中,$\beta = 1.11$,$\mathscr{E}_c = 6.91 \times 10^3$ V cm^{-1}。解这个等式,得 $\mathscr{E}/\mathscr{E}_c = 1.142$ 或者 $\mathscr{E} = 7.89 \times 10^3$ V cm^{-1}。

在结束对强电场下载流子输运的讨论之前,我们对热电子作最后的总结。前面讨论中引入了热电子温度 T_e,描述的是通过碰撞与晶格交换能量的整体电子的特征。某些电子,偶尔会在相当长的时间内不与晶格碰撞,这样获得的速度会超过 v_l,其动能对应的温度也会大于 T_e。尽管这些电子是少数,但它们能量非常高,而且长距离内没有经历散射,所以具有非常重要的物理效应。如果大量的这些"幸运"电子在两次散射间运动的距离与器件尺寸相当时,这些载流子的电荷输运被称为弹道(未散射)输运。

1.2.3 扩散电流

前一节讨论了外电场下遵守欧姆定律的漂移电流。欧姆定律适用于金属和半导体,这在过去的经验中已经熟知。在半导体中,如果材料中存在载流子的能量或浓度的空间变化,那么就会产生另一种重要的电流分量。这一电流分量称为扩散电流。通常扩散电流在金属中并不重要,因为金属具有非常高的电导率。而半导体的电导率较低,载流子能量和浓度的非均匀性使得扩散成为引起电流的一种重要原因。

为了了解扩散电流的起源,考虑一种假想的情况,即一块 n 型半导体只在一维方向上浓度有变化(图 1.19)。假定半导体温度均匀,电子的平均能量不随 x 变化,只有浓度 $n(x)$ 随 x 变化。考虑在平面 $x=0$ 处单位时间在单位面积的电子数。在一定的温度下,电子沿一维 x 方向做无规则热运动,同时假设没有外加电压,在图 1.19 中,电子从左边大约 $x = -\lambda$ 处经过一次碰撞后通过 $x=0$ 平面,(λ 为电子的平均自由程,且 $\lambda = v_{th}\tau_{cn}$)。单位时间从左边通过 $x=0$ 平面单位面积的平均电子数取决于 $x = -\lambda$ 处的电子浓度,为

$$\frac{1}{2}n(-\lambda)v_{th} \tag{1.2.13}$$

因子$\left(\frac{1}{2}\right)$表示 $x = -\lambda$ 处的电子碰撞后将一半向左运动,一半向右运动。同样,单位时间从右边通过 $x=0$ 平面单位面积的电子数为

$$\frac{1}{2}n(\lambda)v_{th} \tag{1.2.14}$$

因此,从左边单位面积流过的粒子的净流量(用符号 F 表示)为

$$F = \frac{1}{2}v_{th}[n(-\lambda) - n(\lambda)] \tag{1.2.15}$$

将 $x = \pm\lambda$ 处的电子浓度用台劳级数展开,取前两项

$$F = \frac{1}{2}v_{th}\left\{\left[n(0) - \frac{dn}{dx}\lambda\right] - \left[n(0) + \frac{dn}{dx}\lambda\right]\right\} = -v_{th}\lambda\frac{dn}{dx} \tag{1.2.16}$$

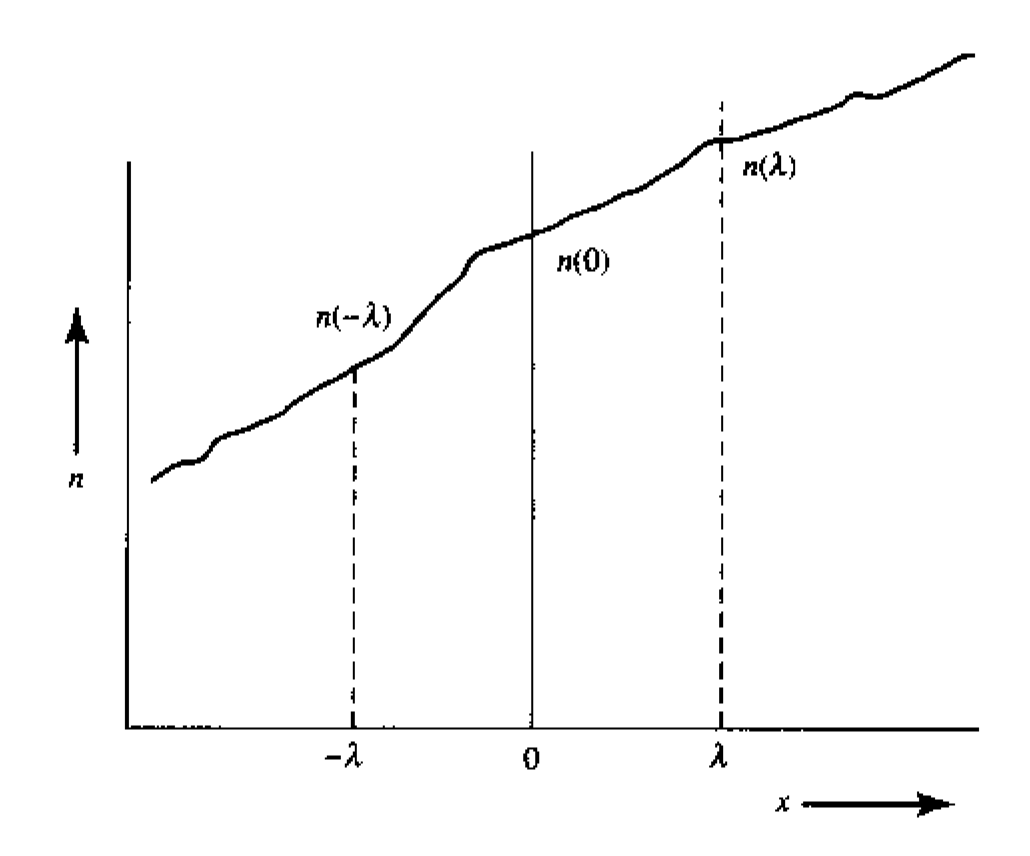


图 1.19　假想的一维固体中电子浓度 n 随距离 x 的变化关系。边界为距原点两侧 λ 的位置，λ 为电子碰撞的平均自由程

由于每个电子的电荷量为 $-q$，粒子流对应的电流为

$$J_n = -qF = q\lambda v_{th}\frac{dn}{dx} \tag{1.2.17}$$

可以看出扩散电流正比于电子的浓度梯度，且起源于存在浓度梯度时带电粒子的随机热运动。如果电子浓度随 x 增加，梯度与电流同为正值。因为电子是从右边的高浓度区流向左边的低浓度区，而电流与电子流动方向相反，所以等式(1.2.17)表示的电流方向在物理上是合理的。

如果将能量均分原理用于一维情况，可将等式(1.2.17)写成更有用的形式

$$\frac{1}{2}m_n^* v_{th}^2 = \frac{1}{2}kT \tag{1.2.18}$$

将关系式 $\lambda = v_{th}\tau_{cn}$ 和等式(1.2.4)代入等式(1.2.17)有

$$J_n = q\left(\frac{kT}{q}\mu_n\right)\frac{dn}{dx} \tag{1.2.19}$$

等式(1.2.19)的右边括号内的量被定义为扩散系数 D_n，即

$$D_n = \left(\frac{kT}{q}\right)\mu_n \tag{1.2.20}$$

等式(1.2.20)被称为 Einstein 关系，它指出固体中描述自由载流子的漂移运动的迁移率与描述载流子扩散运动的扩散系数之间的关系。这里并不是严格的物理推导，考虑固体中的统计力学定律，则可以严格地推导出这个公式。

如果存在外电场，漂移运动和扩散运动将同时存在。那么，总电流为漂移电流和扩散电流之和

$$J_{nx} = q\mu_n n\mathscr{E}_x + qD_n\frac{dn}{dx} \tag{1.2.21}$$

式中，$\mathscr{E}_x$ 为电场在 x 方向上的分量。

空穴也有类似的表达式,总的空穴电流为

$$J_{px} = q\mu_p p\mathscr{E}_x - qD_p\frac{dp}{dx} \tag{1.2.22}$$

因为空穴带正电荷,所以式中取负号。Einstein 关系(等式(1.2.20))也可描述 D_p 和 μ_p 的关系。

上面关于扩散电流的讨论是在载流子浓度不均匀的前提下进行的。这是器件分析中最常见的情况,一般采用等式(1.2.21)和等式(1.2.22)进行分析就足够了。如果半导体中载流子的浓度均匀发布,而某一区域的电子的能量高于另一区域电子的能量,扩散现象也会发生。在这种情况下,必须采用其他公式。习题 1.16 中考虑了载流子具有不同能量时的情况。

例题 扩散电流

已知,硅晶体中平面 x_1(垂直于 x 轴)处的电场不为零,电子在 x 方向上的浓度不均匀,且 x_1 处电子浓度为 10^6cm^{-3}。没有观察到电流流过这一平面。

(a)解释为什么没有电流流动。

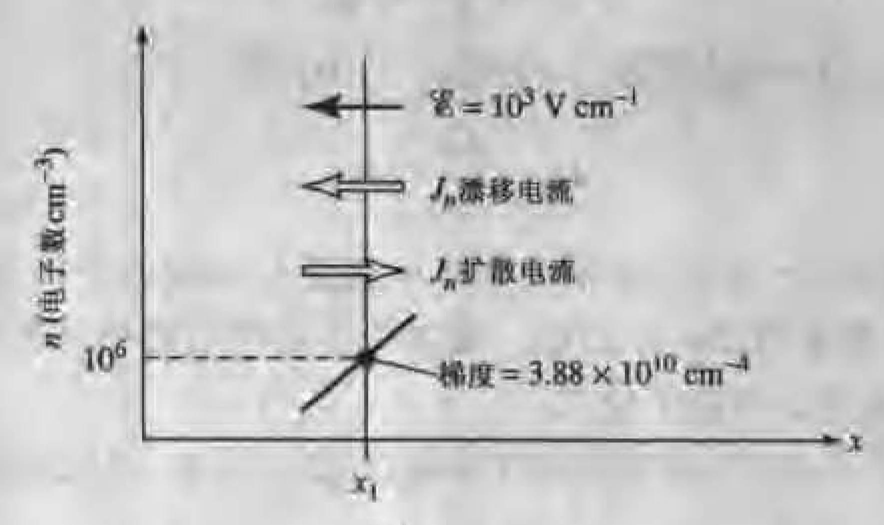
(b)如果场强为 -10^3V cm^{-1}(即沿负 x 方向为 10^3V cm^{-1}),垂直于该平面的电子浓度梯度是多少?

解:(a)根据等式(1.2.5),电场造成的漂移电流为 $J_n = q\mu_n\mathscr{E}$。因为没有电流流动,扩散电流一定等于漂移电流但方向相反。两个分量抵消导致电子电流为零。

(b)根据等式(1.2.21),有

$$J_n = 0 = q\mu_n n\mathscr{E}_x + qD_n\frac{dn}{dx}$$

$$\begin{aligned}\frac{dn}{dx} &= -\frac{\mu_n}{D_n}n\mathscr{E}\\ &= -\frac{q}{kT}n\mathscr{E}\\ &= -\frac{10^6\times(-10^3)}{0.0258}\\ &= 3.88\times10^{10}\text{cm}^{-4}\end{aligned}$$



总电流和准 Fermi 能级 准 Fermi 能级(定义见等式(1.1.28)和等式(1.1.29))在分析处于非平衡态的半导体时非常有用。半导体中有电流流动时就是一个非平衡系统,准 Fermi 能级可以用更简单的形式来定义漂移电流和扩散电流。

为了说明这一点，下面首先推导半导体中用电子能量表示的电场表达式。电场的存在引起带电粒子的能量随位置变化，因此，电子能带由水平（能量为常数）变为倾斜，如图1.14所示。电子能量等于电势乘以电荷 $-q$，因此用本征 Fermi 能级 E_i 表示电场很方便。电场 $\mathscr{E}$ 等于电势的负导数，因此（一维情况下）

$$\mathscr{E} = \frac{1}{q}\frac{dE_i}{dx} = -\frac{d\phi_{fi}}{dx} \tag{1.2.23}$$

式中，准 Fermi 势 $\phi_{fi} = -E_i/q$ 已在等式(1.1.28)中介绍过。电子和空穴浓度可用准 Fermi 能级表示为

$$\begin{aligned} n &= n_i \exp(E_{fn} - E_i)/kT = n_i \exp q(\phi_{fi} - \phi_{fn})/kT \\ p &= n_i \exp(E_i - E_{fp})/kT = n_i \exp q(\phi_{fp} - \phi_{fi})/kT \end{aligned} \tag{1.2.24}$$

将等式(1.2.23)和等式(1.2.24)代入等式(1.2.21)并结合 Einstein 关系（等式(1.2.20)），得到

$$J_n = \mu_n n \frac{dE_{fn}}{dx} = -q\mu_n n \frac{d\phi_{fn}}{dx} \tag{1.2.25}$$

对空穴进行类似的推导，得到

$$J_p = \mu_p p \frac{dE_{fp}}{dx} = -q\mu_p p \frac{d\phi_{fp}}{dx} \tag{1.2.26}$$

等式(1.2.25)和等式(1.2.26)表明，每一种载流子的总电流（漂移和扩散分量之和）正比于该种载流子准 Fermi 能级的梯度 这一简单的表达形式在用能带图形象地描述器件中的总电流时非常有用。这些等式在数学分析中也很有用，它可使许多复杂的问题得到简化。

1.3 器件：霍尔效应磁传感器

按照贯穿本书的结构，下面以讨论一种集成电路器件——霍尔效应磁传感器来结束本章的内容。虽然霍尔效应传感器是非常规集成电路器件，但是却具有重要的商业价值。数以亿计的集成霍尔电路主要作为无接触开关（例如，在计算机终端键盘中）和机械近程探测器而得到应用。霍尔效应磁场检测集成电路，将敏感元件与偏置电路、放大电路和信号处理电路集成在一个硅芯片上，对非电信号具有可控的灵敏度，是集成传感器中非常成功的例子。集成传感器领域发展非常快，因为它是在已经发展得非常精致的纯电学集成电路的基础上进行的，虽然输入信号不是电学量，但也可以很经济地利用现有的电路。例如，微处理器与非电学传感器集成在一起就可构成控制系统中功能强大的元件。除了磁场外，其他非电学输入信号的集成传感器，例如输入信号为可见光、红外辐射、温度、压力、力、加速度、化学气体和湿度等，也都有可观的发展前景。

1.3.1 霍尔效应的物理机制

霍尔效应是1879年由美国物理学家 E. H. Hall 发现，并用他的名字命名的。霍尔效应是带电载流子在磁场中运动时，受到磁场力作用的直接结果。电荷为 q 的粒子，以速度 $\vec{v}$ 在磁感应强度为 $\vec{B}$（两个变量都为矢量）的磁场中运动，受到的作用力为

$$\vec{F} = q\vec{v} \times \vec{B} \tag{1.3.1}$$

式中,矢量叉积(×)表示两个矢量的模相乘再乘以两者间夹角的正弦。

图1.20给出了霍尔效应的示意图。霍尔效应通常用于非本征半导体,所以只有一种载流子起主要作用,另一种载流子的浓度可忽略。为了便于讨论,电子和空穴在图1.20中都被表示出来了。

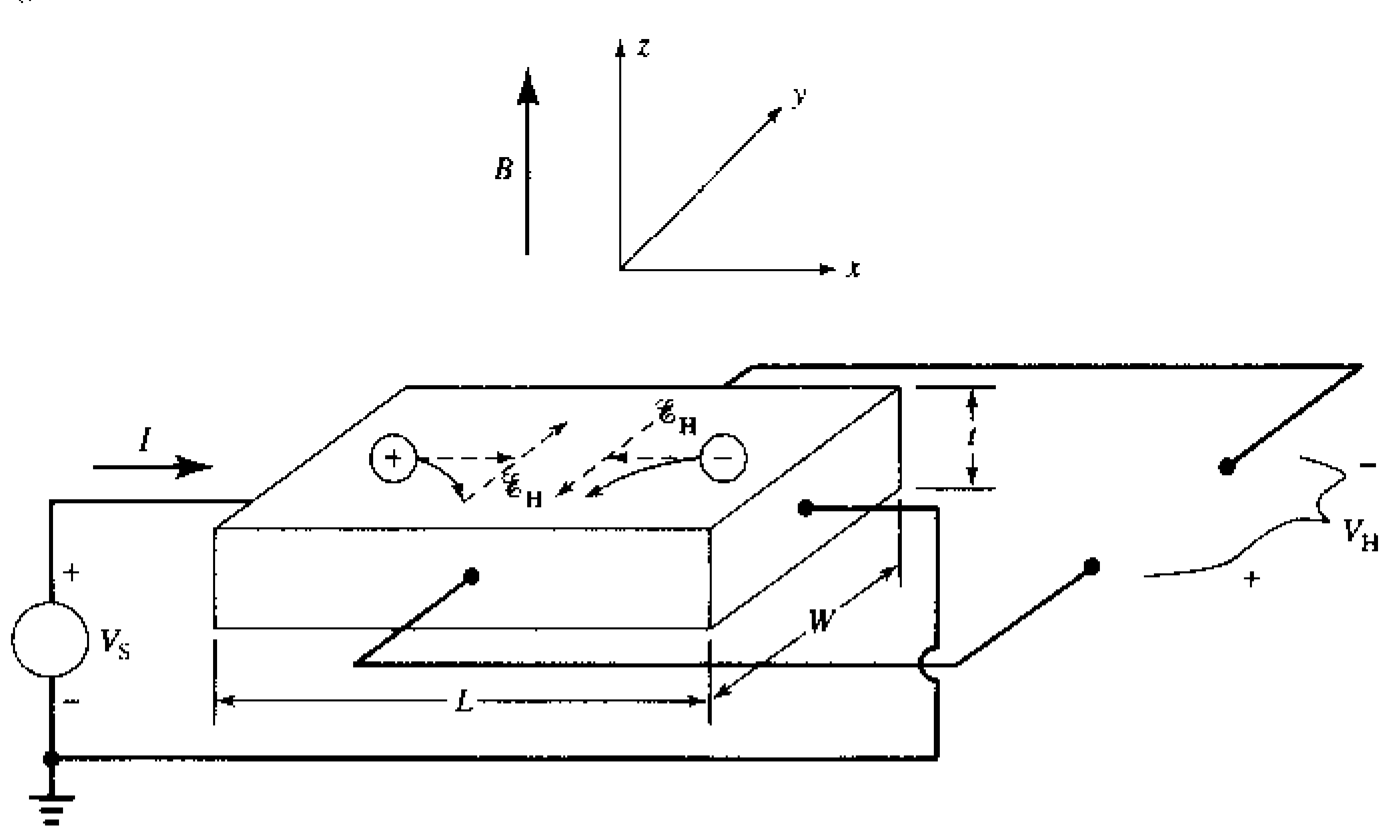


图1.20 霍尔效应示意图。电流沿正 x 方向流动,磁场 B 沿正 z 方向,电子和空穴受磁场力作用沿负 y 方向偏转,其中空穴形成了沿正 y 方向的霍尔电场 $\mathscr{E}_H$,电子形成了沿负 y 方向的霍尔电场 $\mathscr{E}_H$

正如等式(1.3.1)中表示的,半导体中的载流子(电子或空穴)受到的磁场力的方向同时垂直于磁场方向和载流子的运动方向。稳态时,这个作用力与电荷重新分布后形成的电场力相抵消。因为在横向不可能有净的载流子电流,所以这些力必须抵消。产生的电场称为霍尔电场 $\mathscr{E}_H$。将霍尔电场沿半导体的宽度方向积分得到霍尔电压 V_H,这个量也可从半导体两侧的电极测量得到。对均匀半导体和均匀磁场,霍尔电压恰好等于霍尔电场和半导体宽度 W 之积:$V_H = \mathscr{E}_H W$。

在图1.20中,沿正 x 方向的电流由沿 x 正方向流动的空穴和沿相反方向流动的电子组成。因为两种载流子的电荷相反,运动方向也相反,对给定的电流方向,电子和空穴受到的磁场力具有相同的方向。这个作用力引起的电荷偏离从而形成的电子和空穴的霍尔电压具有相反的极性。因此霍尔电压的极性可用来确定半导体材料是n型还是p型。

利用图1.20中表示的各种量可以推导出霍尔效应的基本参量。在图中,电流沿 x 轴正方向,磁场沿 z 轴正方向,因此霍尔电场沿 y 方向。磁场力使电子和空穴都沿负 y 方向偏转,空穴感应的霍尔电场沿正 y 方向,电子的为负 y 方向。考虑载流子的漂移速度 v_d,令磁场力和霍尔电场力相等,有 $q\mathscr{E}_H = qv_dB$。空穴的速度 v_d 与电流的关系为 $v_d = J_x/qp$,电子为 $v_d = -J_x/qn$,这样空穴的霍尔电场可用电流和外加磁场表示为

$$\mathscr{E}_H = \frac{J_xB}{qp} \tag{1.3.2}$$

电子的为

$$\mathscr{E}_{\mathrm{H}} = -\frac{J_x B}{qn} \tag{1.3.3}$$

等式(1.3.2)和等式(1.3.3)可统一表示为

$$\mathscr{E}_H = R_H J_x B \tag{1.3.4}$$

式中,R_{H}为霍尔系数。在这一简单的推导中,空穴的霍尔系数等于$1/qp$,电子的霍尔系数等于$-1/qn$。实际上,如果考虑载流子速度的统计分布,需要对霍尔系数进行修正,这样等式(1.3.4)才可以精确地预测出霍尔电场。考虑这个新的修正因子r,空穴的霍尔系数变为

$$R_{\mathrm{H}} = r/qp \tag{1.3.5}$$

电子的霍尔系数为

$$R_{\mathrm{H}} = -r/qn \tag{1.3.6}$$

因子r的典型值在1~2之间(考虑晶格散射后的理论值为1.18,考虑离化杂质散射后的理论值为1.93)。

霍尔电压V_{H}等于$\mathscr{E}_H$和W的乘积,用总电流I可表示为

$$V_{\mathrm{H}} = \frac{R_{\mathrm{H}} IB}{10^8 t} \tag{1.3.7}$$

式中,R_{H}的单位为$\mathrm{cm^3C^{-1}}$,I的单位为安培,B为Gauss,t为厘米,V_{H}为伏特(将MKS单位制中的米和特斯拉(或韦伯$\mathrm{m^{-2}}$)转换为更常用的半导体单位厘米和Gauss,式中需要因子10^8)。

从等式(1.3.7)可看出在未知半导体中,如果已知磁场和电流,那么,测量霍尔电压可确定出霍尔系数,再利用等式(1.3.5)和等式(1.3.6)可以计算出未知载流子的类型和浓度。从载流子浓度和已知的电流,又可得到材料电导率和霍尔迁移率($\mu_{\mathrm{H}} = (\sigma|R_H|)$。因此,霍尔效应是研究半导体的一项非常有用并常用的实验技术[7]。

1.3.2 集成霍尔效应磁传感器

在集成电路中利用霍尔效应就必须先隔离出一个类似于图1.20中的导电区域。这通常通过外延、氧化物或pn结隔离(工艺过程将在第2章中描述)的方法形成,如图1.21所示。最简单的情况下,在使用霍尔效应理论时,W必须远大于长度L,以使电流密度J在整个样品的截面上是均匀的。实际情况是,如果$L \gg W$,从两端欧姆接触测得的霍尔电压会比较小。在实际生产的集成电路中,W通常与L大小相当。对几何尺寸为矩形的样品,前面推导得到的V_H表达式(等式(1.3.7))可通过乘一个因子K来校正,它近似为1,是宽长比W/L的函数[8]。在前面处理霍尔效应时我们没有考虑这一修正。

在设计集成电路的霍尔效应元件时,器件的功耗是需要考虑的一个重要因素。为了计算功耗,我们用霍尔系数来表示霍尔元件的电阻。例如,对p型元件,可写出电阻R的表达式为

$$R = \frac{\rho L}{A} = \frac{L}{q\mu_p p W t} = \frac{L R_{\mathrm{H}}}{r \mu_p W t} \tag{1.3.8}$$

外加电压为V_{S}时,可写出$I = V_{\mathrm{S}}/R$或

$$I = \frac{rV_S \mu_p W t}{R_H L} \tag{1.3.9}$$

则等式(1.3.7)可表示为

$$V_H = r\mu_p V_S \frac{W}{L} B \times 10^{-8} \tag{1.3.10}$$

与等式(1.3.7)一样,B 以 Gauss 为单位,长度以厘米为单位。

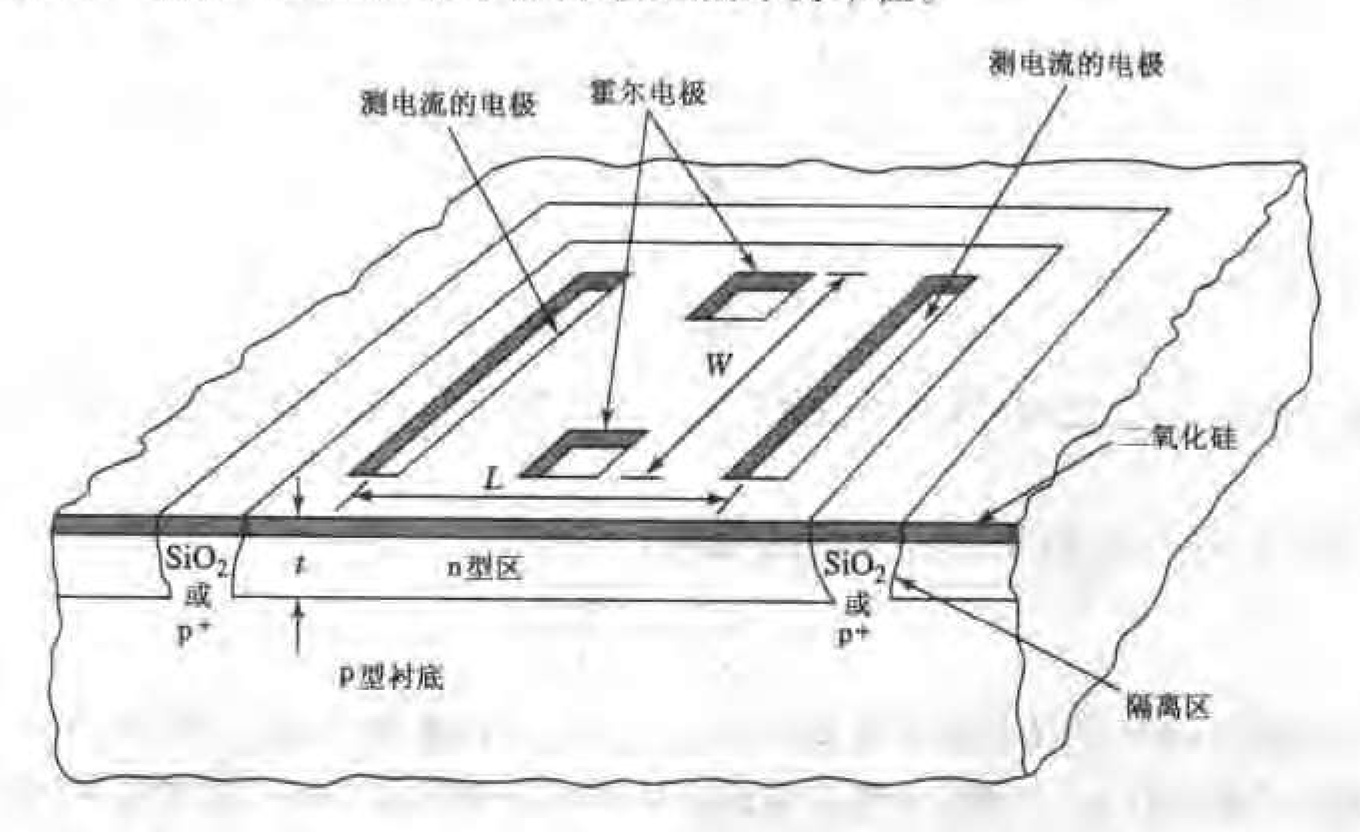


图1.21 用于集成传感器电路的霍尔效应元件。元件制备在高阻 n 型硅上,由第 2 章将要描述的二氧化硅区或 p 型区隔离开。该图表示的是还没有加金属电极的元件

例题 霍尔元件的品质因子

(a)推导霍尔元件的品质因子 M_H,即单位功耗、单位磁场下的霍尔电压。考虑 p 型元件,已知 $R_H = 8 \times 10^3 \text{cm}^3\text{C}^{-1}$,$W/L = 1$,$r = 1.2$,$t = 8\mu\text{m}$。

(b)当 $B = 500$ Gauss,元件功耗为 1.43mW 时,计算元件的电阻和 M_H 值。

解:

(a)根据等式(1.3.10),V_H可用外加电压 V_S表示。功耗 P_H等于 $V_S \times I$。因此

$$P_H = \frac{rV_S^2 \mu_p W t}{R_H L}$$

和

$$M_H = \frac{V_H}{P_H B} = \frac{R_H \times 10^{-8}}{V_S t} = \frac{r \times 10^{-8}}{qpV_S t}$$

单位为 V/Gauss-Watt。从这一推导结果可以看出,通过降低外加电压、减小杂质浓度和霍尔元件的厚度可以提高 M_H。

(b)计算各参量:

1. $p = r/qR_H = 9.38 \times 10^{14} \text{cm}^{-3}$

2. 根据图1.16，$\mu_p = 475 cm^2 V^{-1} s^{-1}$
3. 根据等式(1.3.8)，$R = 17.5 k\Omega$
4. 因为 $P_H = V_S^2 / R$，$V_S = \sqrt{P_H R}$ 或 $V_S = 5V$
5. 因此，$M_H = \dfrac{8 \times 10^3 \times 10^{-8}}{5 \times 8 \times 10^{-4}}$，或 $M_H = 0.02$

在这个例题中可以看出，通过减小杂质浓度可以提高霍尔传感器的性能。当杂质浓度减小时，电子和空穴的迁移率提高(图1.16)，因此给定偏压下的霍尔电压增加(等式(1.3.10))。大部分霍尔传感器是在掺杂浓度约为 $10^{15} cm^{-3}$ 的 n 型硅片上制备的。从图1.16可知在这一浓度范围内，迁移率不随掺杂浓度剧烈变化，因此由于工艺偏差导致的杂质浓度的变化对迁移率的影响很小，对霍尔系数的影响也很小。霍尔元件的典型灵敏度为 30μV/Gauss 量级，而且直到上万 Gauss 时还能表现出极好的线性响应。

集成霍尔电路主要用于检测器件或元件的位置。大部分生产出来的霍尔电路用于无接触键盘开关。在典型的键盘开关应用中，当按下一个键时，柱塞上的永久磁铁发生移动。这一动作在霍尔传感元件附近引入了500Gauss 量级的磁场，产生的霍尔电压大约为15mV，很容易被芯片上的放大器探测到。集成霍尔传感器还用于制备磁场计，这时信号与待测磁场成正比。图1.22给出了一种用于磁场计的商用霍尔效应集成传感器。



图1.22 一种精确的、线性输出的霍尔传感器芯片。集成电路中包括偏置元件，温度补偿电路和片上放大电路。芯片面积为 $1.12 \times 1.98 mm^2$，霍尔元件(右下角大的图形)面积为 $230 \times 335 \mu m^2$(引自 G. B. Hocker, Honeywell Corporation)

小结

固体能带结构是固体电子学的一个重要基础。这个重要概念与孤立原子中的量子化能级有关。金属、半导体和绝缘体电导率的巨大差异就来自这三类材料的能带结构的基本差别。

按能带论,我们可推断有空穴和电子两种类型的载流子存在。大部分情况下,空穴和电子可被看做半导体中经典的自由粒子。如果把自由粒子质量换成空穴或电子的有效质量,它们的运动规律以及按能量的统计分布都可以计算出来。电子和空穴的浓度可通过对纯净的半导体掺杂来控制。在掺杂工艺中,引入三价原子替代晶格中的四价原子硅,则单晶硅中增加了空穴的浓度,这时的晶体称为 p 型硅。为了增加电子浓度得到 n 型硅,则需要掺入五价杂质替代晶格中的硅原子。

用统计物理的原理计算出半导体中电子和空穴的浓度后,就可以了解半导体的许多电学特性。Fermi 能级是描述晶体热平衡时自由载流子浓度的重要参量,通常情况下,半导体中的 Fermi 能级位于不存在电子允许能态的禁带中。这种情况下的半导体称为非简并半导体。如果 Fermi 能级位于能带的允许态中,就称之为简并半导体,其导电特性则与金属类似。当材料处于非平衡态时,需要对电子和空穴分别定义两个不同的准 Fermi 能级。两个准 Fermi 能级间的能量差是半导体偏离热平衡程度的一个度量。当半导体回到热平衡状态时,电子和空穴的准 Fermi 能级又合并为可表征两种载流子浓度的单一 Fermi 能级。在考虑半导体中的光生电子和空穴浓度,以及简化对自由载流子漂移和扩散的表征时,采用准 Fermi 能级很有用。

在半导体电子和空穴的基础理论中,通常假定杂质浓度非常低,不会影响晶体的能带结构。重掺杂时,晶体中的杂质原子彼此之间不再离得很远,它们的存在会影响能带结构。重掺杂对半导体特性最重要的一个影响是本征载流子浓度的有效值 n_{ie} 将随掺杂浓度升高而升高。

半导体中的自由载流子在外加低电场或中等电场作用下,迁移率是常数。电场作用下载流子运动的难易程度用单位电场下的平均漂移速度即迁移率表示。迁移率的大小由自由载流子受到的散射机构决定(即与晶格的作用)。固体中自由载流子因具有热能而随机运动,如果载流子浓度或能量在空间分布不均匀,就会产生净的载流子流。由这一过程导致的载流子的净的流动称为扩散。当晶体温度均匀时,扩散是从载流子浓度高的区域向低的区域运动。对非简并半导体,载流子扩散系数 D 与迁移率 μ 的关系用 Einstein 关系 $D/\mu = kT/q$ 表示。

强电场下,电子和空穴受到的散射主要不再由声子波散射决定,漂移运动的速度也不再与外加电场成正比,这时迁移率不是常数,而是随外加电压的增加而下降的。当外加场强达到 $10^4 \sim 10^5\ \mathrm{V\ cm^{-1}}$ 时,电子和空穴的漂移速度达到其上限,这时必须考虑自由载流子散射的其他模式。

当半导体中有垂直于磁场的电流流动时,运动的带电粒子受到磁场的作用力会引起所谓的霍尔效应。霍尔效应实验为研究半导体中自由载流子的特性提供了有力的工具。制作在单个硅片上,由霍尔元件、偏置电路、温度补偿电路和放大器组成的集成传感器电路已经是重要的商业产品了。

参考文献

1. (a) F. Morin and J. P. Maita, *Phys. Rev.* **96**, 28 (1954). (b) A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, New York (1967).
2. P. P. Debye and E. M. Conwell, *Phys. Rev.* **93**, 693 (1954).
3. (a) D. M. Caughey and R. F. Thomas, *Proc. IEEE* **55**, 2192 (December 1967). (b) G. Masetti, M. Severi, and S. Solmi, *IEEE Trans. Electr. Devices* **ED-30**, 764 (July 1983).
4. N. D. Arora, J. R. Hauser, and D. J. Roulston, *IEEE Trans. Electr. Devices* **ED-29**, 292 (February 1982).
5. C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, *Solid-State Electronics* **20**, 77 (February 1977).
6. J. L. Moll, *Physics of Semiconductors*, McGraw-Hill,

New York (1964), p.198.

7. E. H. PUTLEY, *The Hall Effect and Related Phenomena*, Butterworths, London (1960).
8. J. T. MAUPIN and M. L. GESKE, *The Hall Effect in Silicon Circuits*, Symposium on the Hall Effect, Johns Hopkins University, Baltimore, Md. (1981), p. 421.
9. S. WANG, *Solid-State Electronics*, McGraw-Hill, New York (1966), p. 263.
10. S. M SZE and J. C. IRVIN, *Solid-State Electronics* 11, 599 (1968).
11. J. L. MOLL, *Physics of Semiconductors*, McGraw-Hill, New York (1964), p 99.
12. R B ADLER, A. C SMITH, and R. L. LONGINI, *SEEC*, Vol. 1, *Introduction to Semiconductor Physics*, Wiley, New York (1964).

参考书

B. G. STREETMAN and S. BANERJEE *Solid-State Electronic Devices*, 5th ed. Prentice-Hall, Simon and Schuster Co., N.J. (2000).

R. F. PIERRET, *Semiconductor Fundamentals*, Vol. I, *Modular Series on Solid-State Devices*, Addison-Wesley, Reading, Mass. (1983).

J. P. MCKELVEY, *Solid-State and Semiconductor Physics*. Original publisher Harper & Row, New York (1966), reprinted by Dover Publications, New York (1984).

W. E. BEADLE, J. C. C. TSAI, and R. D. PLUMMER, Editors, *Quick Reference Manual for Silicon Integrated-Circuit Technology*, Wiley-Interscience, New York (1985) (for silicon properties).

习题

1.1 纯硅样品中掺入浓度为$10^{16}cm^{-3}$的施主杂质P。假定P原子在硅中均匀分布，且已知P的原子量为31。

(a)计算300K时样品的电阻率；

(b)计算施主杂质在晶体中的质量比；

(c)如果除P原子外，再掺入$10^{17}cm^{-3}$的B原子且均匀分布，请问最终的电阻率和导电类型(即p型还是n型材料)；

(d)画出(c)问的能带图，并标出Fermi能级的位置。

1.2* 已知硅中的浅施主浓度如下：

(a)$1\times10^{16}cm^{-3}$的B原子

(b)$3\times10^{16}cm^{-3}$的As原子和$2.9\times10^{16}cm^{-3}$的B原子

计算27℃时平衡态的电子和空穴浓度及Fermi能级的位置。

1.3 已知n型硅样品均匀掺有浓度为$N_d=10^{16}cm^{-3}$的As原子，p型硅样品掺有浓度为$N_a=10^{15}cm^{-3}$的B原子。对每一种半导体材料计算：

(a)假定所有的可动电子和可动空穴来自杂质原子，如果只有一半杂质原子离化，求相应的温度(提示：利用等式(1.1.21)和等式(1.1.22)，并且考虑$f_D(E_f)=1/2$)；

(b)如果本征载流子浓度n_i是杂质浓度的10倍，求相应的温度，参见表1.4中的$n_i(T)$；

(c)假定所有杂质全部离化，求300K时的少子浓度；

(d)求300K时两种材料的Fermi能级相对于价带边E_v的距离。如果两种类型的杂质存在于同一个样品中，计算Fermi能级的位置。

1.4* 已知27℃时n型硅样品的电阻率为5Ω·cm。在热平衡条件下，计算27℃，100℃和500℃时的空穴浓度(参考相关图表查出所需要的参数)。

1.5* 晶粒间界和其他结构缺陷在多晶硅禁带中会引入深能级的允许态。假设每一缺陷引入两个分立的能级：价带顶以上0.51eV处的受主能级和价带顶以上0.27eV处的施主能级(注意$E_a>E_d$，且都不是浅杂质能级)。已知缺陷态每一种带电状态(+、-或中性)在热平衡时的数值比由以下关系给出(见参考文献11)

$$N_d^+ : N_0 : N_a = \exp\frac{E_d - E_f}{kT} : 1 : \exp\frac{E_f - E_a}{kT}$$

(a)当 Fermi 能级从 E_v 向 E_c 移动时，画出这三类缺陷浓度的变化图。哪种缺陷在重掺杂 p 型材料中占主导地位？在 n 型材料中呢？

(b)缺陷对多数载流子浓度的影响如何？

(c)利用上面的信息，确定缺陷能级的带电状态和在硅晶体中不含浅杂质原子时 Fermi 能级的位置。该样品是 n 型还是 p 型？

(d)当样品中掺有 $2\times10^{17}\text{cm}^{-3}$ 的 P 原子，并有 $5\times10^{16}\text{cm}^{-3}$ 的缺陷时，计算电子和空穴的浓度及 Fermi 能级的位置？

1.6* 已知一块半导体中有两种主要的散射机制。如果只有第一种散射机制存在，迁移率为 $800\text{cm}^2(\text{Vs})^{-1}$。如果只有第二种散射机制存在，迁移率为 $200\text{cm}^2(\text{Vs})^{-1}$。在同时考虑两种散射机制时迁移率为多少？

1.7 计算电阻率为 $2.8\times10^{6}\Omega\cdot\text{cm}$、密度为 2.7g cm^{-3} 的 Al 中的电子迁移率。已知 Al 的原子量为 27，Al 的三个价电子中，室温下平均有 0.9 个电子是自由的、可参与导电。如果取 $m^* = m_0$，计算两次碰撞之间的平均时间，并将此值与轻掺杂硅所对应的值比较。

1.8* 已知 27℃时轻掺杂硅中电子在外加电场作用下运动，漂移速度为热运动速度的十分之一。计算电子漂移 1μm 长度时遭受的平均碰撞次数。加在这一区域的外加电压为多少？

1.9 已知一块均匀的长 2μm，电阻率为 $1\Omega\cdot\text{cm}$ 的 n 型硅材料，加在其两端的电压增加一倍时，电流只增加了 50%，请解释原因(在两种电流下材料均保持电中性)

1.10* 室温下，在一块均匀的轻掺杂 n 型硅中，电子浓度从 $x=0$ 处的 10^{17}cm^{-3} 线性变化到 $x=2\mu\text{m}$ 处的 $6\times10^{16}\text{cm}^{-3}$，且不随时间变化。如果没有外加电场时，计算硅中的电子电流密度。假定 $\mu_n = 1000\text{cm}^2\text{V}^{-1}\text{s}^{-1}$，$T = 300\text{K}$。

1.11+ 如果在 GaAs 中掺入硅原子，则硅可以代替三价的镓原子或五价的砷原子，假设硅原子完全离化，且 10^{10}cm^{-3} 的硅原子中有 5% 替代 Ga 原子，95% 替代 As 原子，当样品温度为 300K 时

(a)计算施主浓度和受主浓度；

(b)计算电子浓度和空穴浓度及 Fermi 能级的位置；

(c)假设晶格散射机制占主导地位，计算 GaAs 的电导率；

参见表 1.3 查找 GaAs 的相关参数。

1.12+ (固体中的介质弛豫)考虑一块均匀的只含有一种载流子的导体，其电导率为 σ、介电常数为 ϵ。已知 $t=0$ 时刻，可动电荷在空间的浓度分布为 $\rho(x,y,z;t=0)$。如果忽略扩散电流，根据电磁理论的知识有

$$\nabla\cdot D = \rho;\quad D = \epsilon\mathscr{E};\quad J = \sigma\mathscr{E};\quad \nabla\cdot J = \frac{-\mathrm{d}\rho}{\mathrm{d}t}$$

(a)根据以上条件证明：$\rho(x,y,z;t) = \rho(x,y,z;t=0)\exp[-t/(\epsilon/\sigma)]$，这一结果表明未中和的电荷不能在均匀的导体材料中存在，只能在不连续的界面，表面或其他非均匀的位置积累。

(b)在以下条件下计算介电弛豫时间 ϵ/σ：本征硅；施主浓度为 10^{16}cm^{-3} 的硅；$\sigma = 10^{-16}(\Omega\text{-cm})^{-1}$ 的热氧化 SiO_2[12]。

1.13+ 自由载流子具有热能，所以可在晶格中不断地运动。热平衡时通过任意平面的所有载流子的净流量为零，那么考虑相互抵消的各分量非常有用。这些分量具有重要的物理意义，根据这些分量可以计算出扩散电流的大小。例如，在一给定方向上阻止所有的载流子的流动，就会破坏热平衡条件，这是有重要意义的。已知 $J_x = -qn_0v_x$，证明在固体中任意方向，热运动形成的电流为

$$J = \frac{-qn_0v_{th}}{4}$$

式中 v_{th} 为平均热运动速度，n_0 为自由电子浓度(提示，考虑通过固体的电子流分布在 2π 立体角中)[12]

1.14* 在本征锗、硅、砷化镓和 SiO_2 中，计算产生电子-空穴对所需要的激发光的波长，并确定每种情况

下所处的光谱范围(例如,红外、可见、紫外和X射线)

1.15[*] 对某一简并材料(即必须考虑 Fermi-Dirac 统计,因为 Fermi 能级可能进入允许带内),D 和 μ 之间的关系为

$$\frac{D}{\mu} = \frac{1}{q}\frac{\mathrm{d}E_f}{\mathrm{d}(\ln n)}$$

证明:当 Boltzmann 统计成立,即材料为非简并时,这一关系式可简化为 Einstein 关系 $D/\mu = kT/q$。

1.16 热探针装置是一种有用的实验室设备,它可以用来确定半导体样品的导电类型,由两个探针和一个可指示电流方向的电流计组成。其中一个探针被加热(最简单的情况是用电烙铁尖),另一个探针处于室温下。即使不加外电压,当探针接触到半导体时也会有电流流过。考虑扩散电流的作用,解释这一装置的工作原理,并画图说明 p 型和 n 型半导体的电流方向。

1.17 考虑一维情况下的简单扩散模型。假定所有粒子在相等时间间隔内不连续地运动。运动时,每个粒子每一步只能在空间中以相同的几率向左或向右跳动一个单位。在 $x=0$ 处,$t=0$ 时刻有 1024 个粒子,依次确定经过 10 个时间间隔后粒子在空间的分布。画出每一步跳跃后粒子在空间的分布图,并计算最大浓度的一半处所对应的点之间的距离 W。然后画出 W^2 随时间的变化。对扩散过程中粒子的扩散速率你能说明什么问题(至少一维情况)?[12]

1.18 从以下讨论可以看到表 1.3 和表 1.4 中的值并不完全一致。

(a) $T=300\mathrm{K}$ 时用表 1.4 中的公式计算 E_g 并与表 1.3 中的结果比较。

(b) 用等式(1.1.25)和表 1.3 中的 N_c 和 N_v 值及两个 E_g 值分别计算其 n_i。

(c) 用表 1.4 中的温度变化公式计算 300K 时的 n_i。

[计算值的不一致性反映出这样的事实,即表中的值是由多种实验方法测量得到的。根据许多实验结果,300K 时硅的 n_i 值一般认为是 $1.45\times10^{10}\mathrm{cm}^{-3}$。但许多参量的精度并不知道。]

1.19[*] 在近乎本征的材料中,电子和空穴的运动对霍尔效应都很重要。证明电子和空穴同时存在时霍尔系数为

$$R_H = \frac{\mathscr{E}_H}{J_x B} = \frac{\mu_p^2 p - \mu_n^2 n}{q(\mu_p p + \mu_n n)^2}$$

证明这一结果与 1.3 节中推导出的简单理论是一致的。

本章附录 电场分布、电荷分布和 Gauss 定理

从 1.2 节可知,固体中的外加电场使得电子和空穴运动,产生漂移电流。如果我们知道电场的大小和方向,就可以根据材料特性估算漂移电流。这只是反映固体中了解电场知识重要性的一个例子。尽管有关电场的分析已在普通物理课程中讲授过,鉴于它对器件电子学有着非常重要的应用,这里我们再简单回顾一下。需要进一步了解这些知识的读者,可以在电磁场理论的书籍中查看更详细的讨论。

对器件分析,最重要的是电荷和电场间的关系式。一维情况下(器件分析中最常用),关于电荷密度 $\rho(x)$ 和电场 $\mathscr{E}$ 的微分方程为

$$\mathrm{d}(\epsilon\mathscr{E})/\mathrm{d}x = \rho \qquad (1\mathrm{A}.1)$$

在等式(1A.1)中,$\epsilon=\epsilon_r\epsilon_0$ 为材料的介电常数,等于相对介电常数 ϵ_r 和真空介电常数 ϵ_0 的乘积。

著名的英国科学家 Michael Faraday(1791 到 ~1867),通过引入作用在电荷上的从正电荷出发终止于负电荷的电力线的概念,对等式(1A.1)给出了非常有用的解释。电力线总是与电场平行的,在某一区域内电力线的密度正比于这一区域中电场的大小。Faraday 的关于电力线的概念给出了一个直接证明 Gauss 定理的方法。Gauss 定理指出,将电场在一个封闭的表面积分等于包含在这个表面内的净电荷除以介电常数。

为了使这一概念在器件电子学中变为有用的形式,可以对等式(1A.1)沿 x 方向积分得到一维情况下的

Gauss 定理(通常情况下 ϵ 不是 x 的函数)

$$\epsilon(\mathscr{E}_2 - \mathscr{E}_1) = \int_{x_1}^{x_2} \rho(x)\mathrm{d}x \tag{1A.2}$$

用 Faraday 关于电力线的概念很容易解释等式(1A.2),因为它从数学上描述了在边界上(x_1 和 x_2 之间)介电常数与电场的乘积之差等于在这一区域内总的电荷。正电荷使电力线增加,而负电荷使电力线减小;因此,在两个边界处电场的变化乘以介电常数等于这一区域中两个边界间的电荷总和。一种特殊的情况是电荷 σ 均匀分布在一个有限体积内的表平面上(即在 $x=x_0$ 处的电荷为 σ),根据等式(1A.2),在 $x=x_0$ 时处的电场将发生突变,电场变化值 $\Delta\mathscr{E}$ 为 σ/ϵ。从等式(1A.2)中看不出介电常数发生变化造成的影响,例如介电常数在两种不同材料的边界处会发生改变。因为材料的介电常数表示电场对其分子的极化效应,它的变化导致该边界处会产生有效表面电荷。结果,假如在 $x=x_1$ 处的介电常数从 ϵ_a 变到 ϵ_b,该处的电场也必须从 a 区中的 $\mathscr{E}_a$ 值变到 b 区中的 $\mathscr{E}_b$ 值,使得 $\epsilon_a\mathscr{E}_a = \epsilon_b\mathscr{E}_b$。

在电场 $\mathscr{E}$ 中,作用在电子上的力 F 等于电荷 $-q$ 与电场的乘积

$$F = -q\mathscr{E} \tag{1A.3}$$

电荷移动距离 $\mathrm{d}x$ 所需要做的功 $\mathrm{d}W$ 为

$$\mathrm{d}W = -q\mathscr{E}\mathrm{d}x = q\mathrm{d}\phi \tag{1A.4}$$

式中,电场 $\mathscr{E}$ 与电势 ϕ 的关系为

$$\mathscr{E} = -\mathrm{d}\phi/\mathrm{d}x \tag{1A.5}$$

将等式(1A.5)与等式(1A.1)结合在一起得到 Poisson 方程

$$\mathrm{d}^2\phi/\mathrm{d}x^2 = -\rho/\epsilon \tag{1A.6}$$

等式(1A.6)中已经假定介电常数 ϵ 与 x 无关。

等式(1A.6)是关于 ϕ 的二阶线性微分方程,因此它的解中有两个常数需要用边界条件来确定。因为等式(1A.6)是线性的,它的解具有叠加性,电荷密度 ρ_1 对应的电势的解 ϕ_1 可以与电荷密度 ρ_2 对应的电势的解 ϕ_2 相加,即电荷密度 $\rho=\rho_1+\rho_2$ 对应的电势为 $\phi=\phi_1+\phi_2$。电势解的叠加性被称为叠加原理,通常在解方程时有辅助作用。

例题 用 Gauss 定理求电容

(a)已知一个由两块面积为 A、间距为 W、中间为空气的导体平板组成的平行板电容器。两平行板间所加电压为 V_0。画出两平行板之间($0 \leqslant x \leqslant W$)的(i)电场分布,(ii)电荷分布,(iii)电势分布。

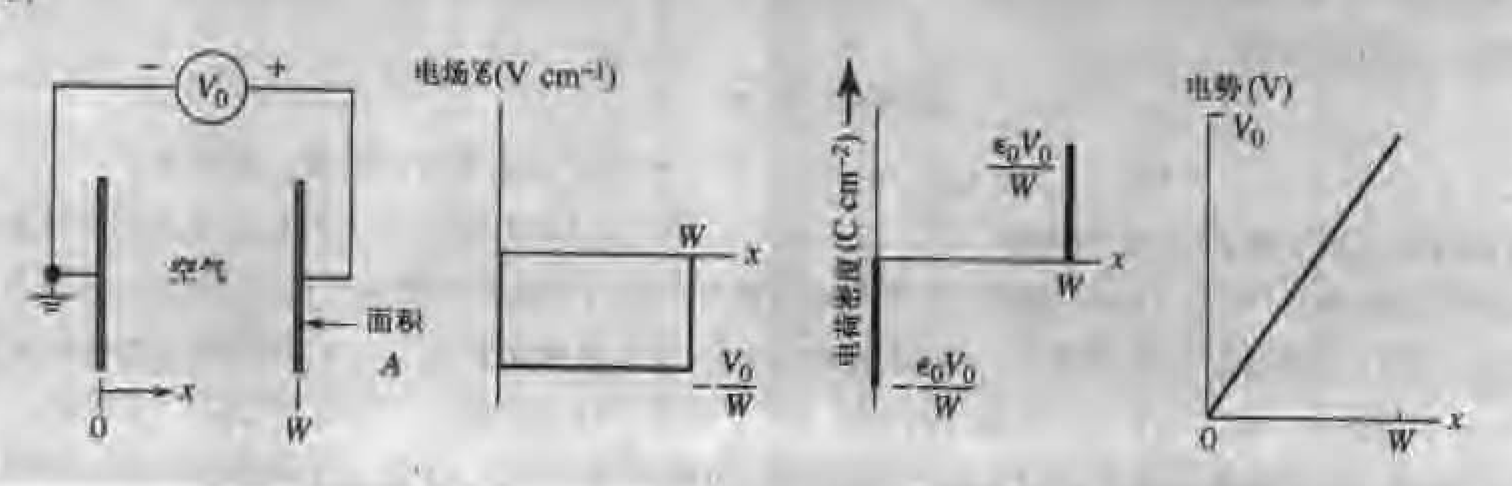
(b)一块没有与电源相连的面积为 A、厚度为 $W/2$ 的理想导体平板,插入平行板电容器中间。重复计算(a)中的问题。

(c)在(b)问中如果插入的不是导体平板,而是 $\epsilon_r=2$,宽度为 $W/2$ 的介质板,重复(a)问中的问题。

(d)这一例题的(a),(b)和(c)中测量到的是两个导体平板之间的什么电容?

解:

(a)



(b)设空气中的电场为$\mathscr{E}_a$，介质中的电场为$\mathscr{E}_d$。

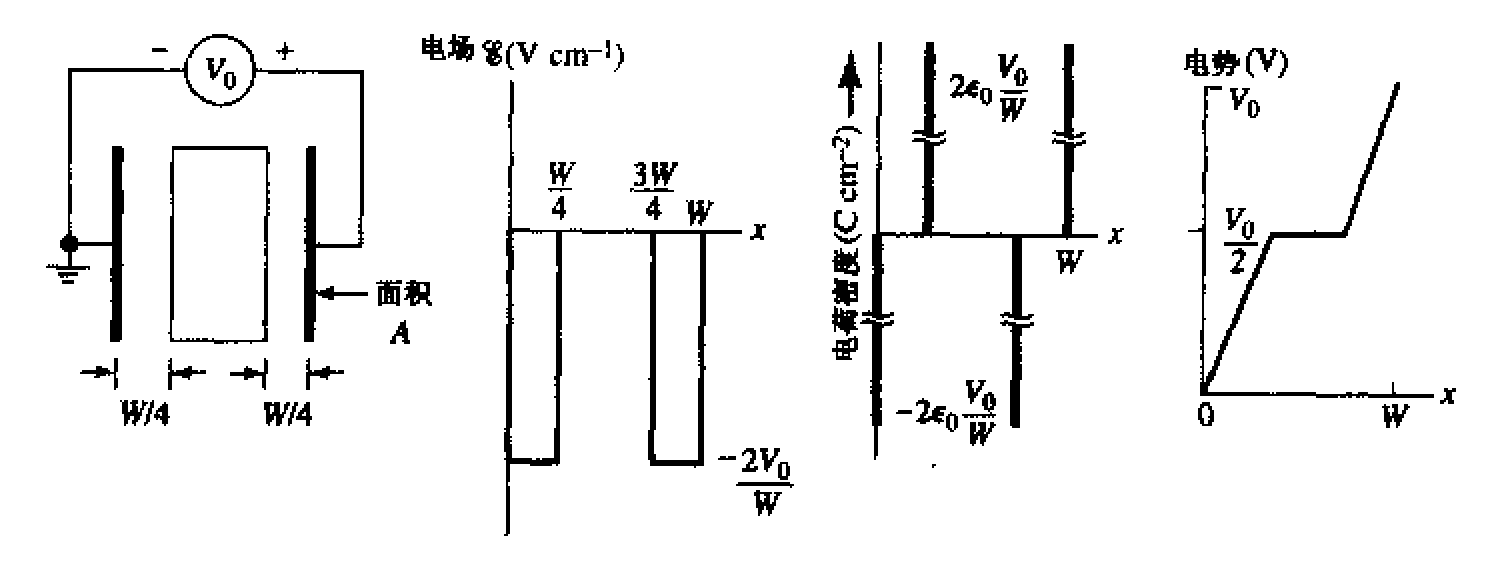
在边界($x = W/4$)处，有 $\epsilon_0 \mathscr{E}_a = \epsilon_r \epsilon_0 \mathscr{E}_d$，因此，

$$\mathscr{E}_d = \mathscr{E}_a/\epsilon_r = \mathscr{E}_a/2$$

平板间的电压等于 V_0，因此

$$\mathscr{E}_a\left(\frac{W}{4}\right) + \frac{\mathscr{E}_a}{2}\left(\frac{W}{2}\right) + \mathscr{E}_a\left(\frac{W}{4}\right) = V_0 \quad 和 \quad \mathscr{E}_a = \frac{4V_0}{3W}$$

平板上的电荷 $= \epsilon_0 \mathscr{E}_a = \epsilon_0 \dfrac{4V_0}{3W}$



(c)

(d)电容 $C = Q/V$，对于(a)，$Q = \epsilon_0 V_0/W$，$C = \epsilon_0/W$；

对于(b)，$Q = 2\epsilon_0 V_0/W$，$C = 2\epsilon_0/W$；

对于(c)，$Q = \frac{4}{3}\epsilon_0 V_0/W$，$C = \frac{4}{3}\epsilon_0/W$。

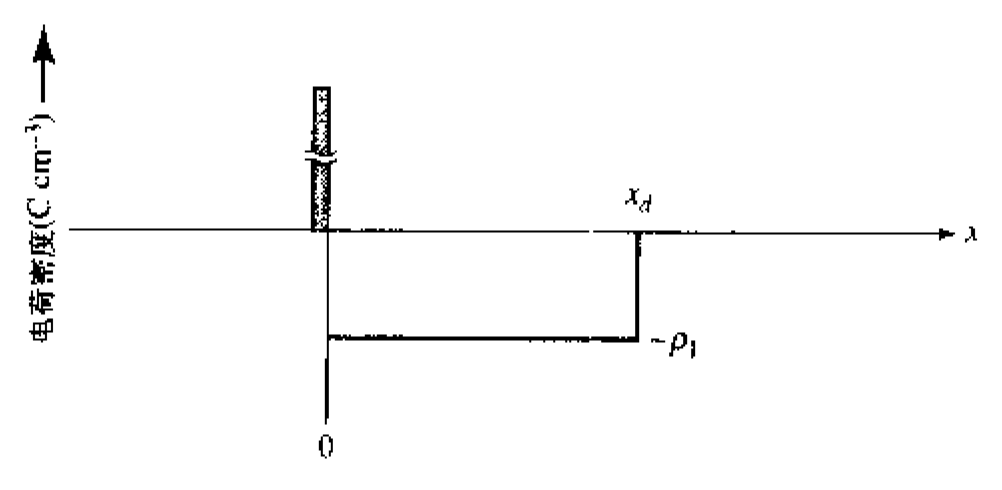
附录习题

A1.1　下图是理想的金属-半导体接触处的电荷分布。在 $x = 0$ 处，面密度为 $Q' = \rho_1 \times x_d$ 的一薄层的正电荷(Δ 函数)与半导体中均匀分布的负电荷相抵消以保持电中性。半导体的介电常数为 ϵ_s。

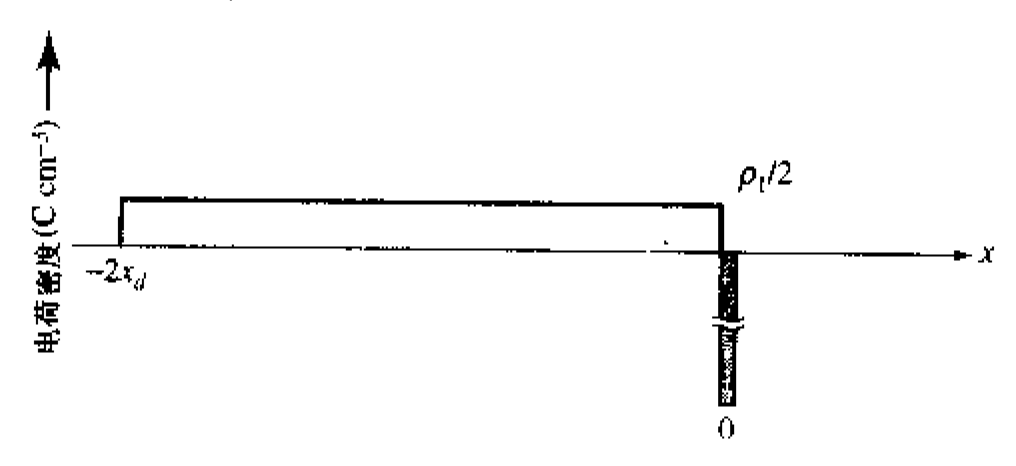
(a)推导$\mathscr{E}(x)$在 $-\infty < x < +\infty$ 之间的表达式，并画出$\mathscr{E}(x)$的分布。

(b)推导 $\phi(x)$ 在 $-\infty < x < +\infty$ 之间的表达式，并画出 $\phi(x)$ 的分布。

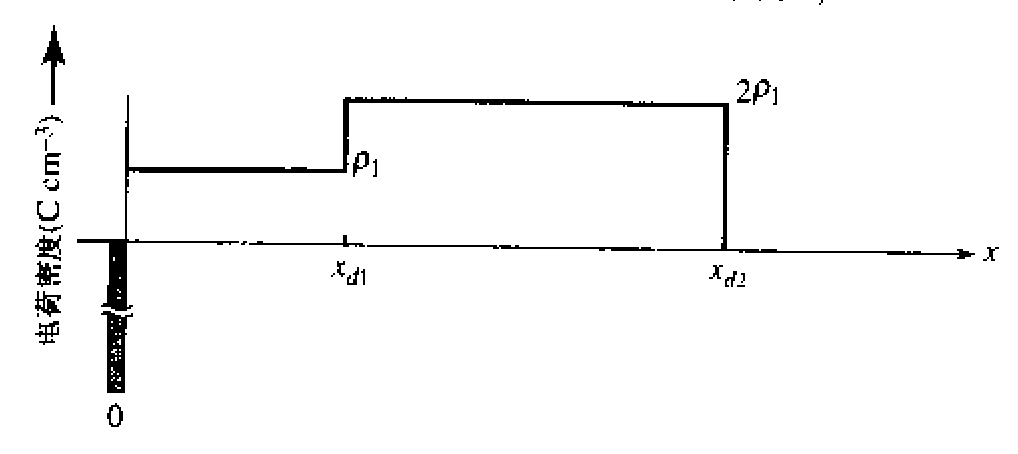
(c)计算 $x = 0$ 和 $x = x_d$ 之间的电位差 $\Delta\phi$(即 $\phi(0) - \phi(x_d)$)。



(d)已知电荷分布如下图时,画出所对应的$\mathscr{E}(x)$和$\phi(x)$ 直觉会告诉你结果是怎样的,而不必重复前面的计算 所有的图必须按适当的比例绘制。



A1.2 下面的金属-半导体接触的电荷分布图中,已知,半导体中有两个电荷密度不同的区域. 在$x=0$处的一薄层负电荷与半导体中总的正电荷相等 半导体中的介电常数为ϵ_s。

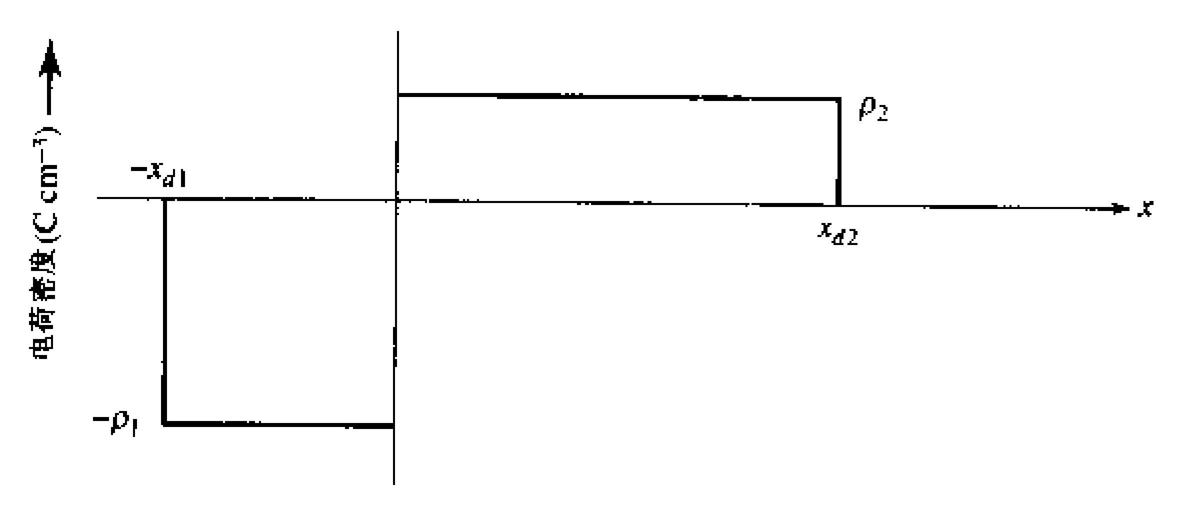


(a)推导$\mathscr{E}(x)$在$-\infty<x<+\infty$之间的表达式,并画出$\mathscr{E}(x)$的分布;

(b)推导$\phi(x)$在$-\infty<x<+\infty$之间的表达式,并画出$\phi(x)$的分布.

(c)计算$x=0$和$x=x_{d1}$之间的电位差$\Delta\phi_1$(即$\phi(0)-\phi(x_{d1})$)和$x=0$和$x=x_{d2}$之间的电位差$\Delta\phi_2$(即$\phi(0)-\phi(x_{d2})$)

A1.3 下面所示的电荷分布与半导体的pn结有关。n型半导体中的总的正电荷与p型半导体中总的负电荷相等(即$\rho_2\times x_{d2}=\rho_1\times x_{d1}$). 半导体中两个区域的介电常数均为$\epsilon$。



(a)推导$\mathscr{E}(x)$在$-\infty < x < +\infty$之间的表达式，并画出$\mathscr{E}(x)$的分布。

(b)推导$\phi(x)$在$-\infty < x < +\infty$之间的表达式，并画出$\phi(x)$的分布

(c)计算$x = x_{d2}$和$x = -x_{d1}$之间的电位差$\Delta\phi$(即$\phi(x_{d2}) - \phi(-x_{d1})$)

(d)用叠加原理和问题A1.1的结果解$\mathscr{E}(x)$和$\phi(x)$。

A1.4 下图给出了金属-氧化物-半导体(MOS)结构的电荷分布。氧化物的介电常数为ϵ_{ox}，半导体的介电常数为ϵ_s。金属中在$x=0$处有一薄层电荷与半导体中浓度为ρ、厚度为x_d的总电荷保持平衡。注意在氧化层内部，$\rho(x) = 0$

(a)推导$\mathscr{E}(x)$在$-\infty < x < +\infty$之间的表达式，并画出$\mathscr{E}(x)$的分布。

(b)推导$\phi(x)$在$-\infty < x < +\infty$之间的表达式，并画出$\phi(x)$的分布

(c)计算电位差$\Delta\phi_0$(即$\phi(0) - \phi(x_0)$)。这是加在氧化物上的电压

(d)计算$\Delta\phi_d$(即$\phi(x_0) - \phi(x_0 + x_d)$)。这是加在半导体上的电压。

(e)用叠加原理和前面的结果分解电荷，解出$\mathscr{E}(x)$和$\phi(x)$(用一种易于利用前面结果的方法来分解电荷)

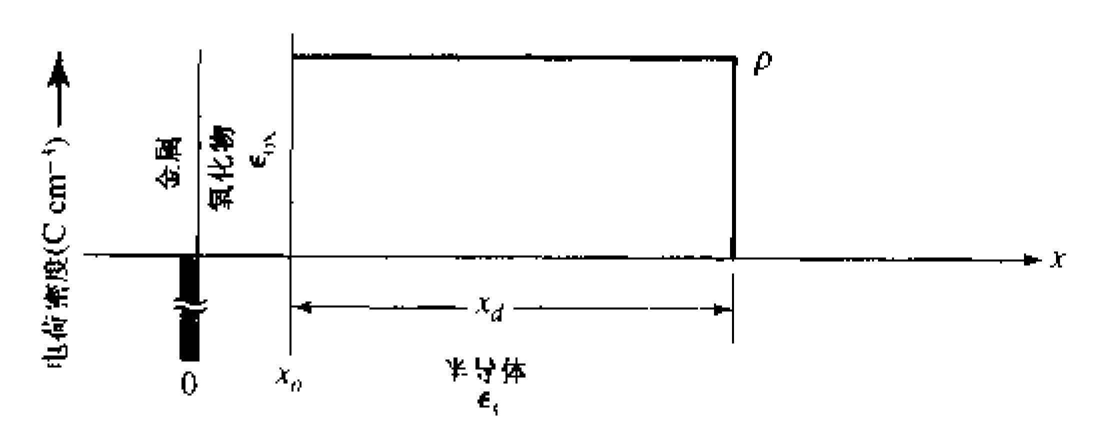


表1.3 半导体和绝缘体的性质(除特别指出，均为300K)

性质	符号	单位	Si	Ge	GaAs	GaP	SiO_2	Si_3N_4
晶格结构			金刚石	金刚石	闪锌矿	闪锌矿	大部分IC工艺中是无定形结构	
元胞中的原子数			8	8	8	8		
原子序数	Z		14	32	31/33	31/15	14/8	14/7
原子或分子量	MW	g/g-mole	28.09	72.59	144.64	100.70	60.08	140.28
晶格常数	a_0	nm	0.54307	0.56575	0.56532	0.54505		0.775
原子或分子浓度	N_0	cm^{-3}	5.00×10^{22}	4.42×10^{22}	2.21×10^{22}	2.47×10^{22}	2.20×10^{22}	1.48×10^{22}
密度		$g\ cm^{-3}$	2.328	5.323	5.316	4.13	2.19	3.44
带隙宽度 300K	E_g	eV	1.124	0.67	1.42	2.24	~8-9	4.7
0K	E_g	eV	1.170	0.744	1.52	2.40		
与温度的依赖关系	$\Delta E_g/\Delta T$	$eV\ K^{-1}$	-2.7×10^{-4}	-3.7×10^{-4}	-5.0×10^{-4}	-5.4×10^{-4}		
相对介电常数	ϵ_r		11.7	16.0	13.1	10.2	3.9	7.5
折射率	n		3.44	3.97	3.3	3.3	1.46	2.0
熔点	T_m	℃	1412	937	1237	1467	~1700	~1900
蒸气压		Torr(mm Hg)(at℃)	10^{-7}(1050) 10^{-5}(1250)	10^{-9}(750) 10^{-7}(880)	1(1050) 100(1220)	10^{-6}(770) 10^{-4}(920)		
比热	C_p	$J(g\ K)^{-1}$	0.70	0.32	0.35		1.4	0.17
热导率	κ	$W(cm\ K)^{-1}$	1.412	0.606	0.455	0.97	0.014	0.185(?)
热扩散系数	D_{th}	$cm^2 s^{-1}$	0.87	0.36	0.44		0.004	0.32(?)
线性热膨胀系数	α'	K^{-1}	2.5×10^{-6}	5.7×10^{-6}	5.9×10^{-6}	5.3×10^{-6}	5×10^{-7}	2.8×10^{-6}
本征载流子浓度*	n_i	cm^{-3}	1.45×10^{10}	2.4×10^{13}	9.0×10^{6}			
电子迁移率	μ_n	$cm^2(V\ s)^{-1}$	1417	3900	8800	300	20	

续表

性质	符号	单位	Si	Ge	GaAs	GaP	SiO_2	Si_3N_4
空穴迁移率	μ_p	$cm^2(V \cdot s)^{-1}$	471	1900	400	100	$\sim 10^{-8}$	
导带等效态密度	N_c	cm^{-3}	2.8×10^{19}	1.04×10^{19}	4.7×10^{17}			
价带等效态密度	N_v	cm^{-3}	1.04×10^{19}	6.0×10^{18}	7.0×10^{18}			
击穿电场强度	$\mathscr{E}_1$	$V\ cm^{-1}$	3×10^5	8×10^4	3.5×10^5		$6-9\times10^6$	
电子有效质量	m_n^*/m_0		1.08^a 0.26^b	0.55^a 0.12^b	0.068	0.5		
空穴有效质量	m_p^*/m_0		0.81^a 0.386^b	0.3	0.5	0.5		
电子亲和势	qX	eV	4.05	4.00	4.07	~4.3	1.0	
每个声子散射失去的平均能量		eV	0.063	0.037	0.035			
电子的光学声子平均自由程	λ_{ph}	nm	6.2	6.5	3.5			
空穴的光学声子平均自由程	λ_p	nm	4.5	6.5	3.5			

来源:A. S. Grove, Physics and Technology of Semiconductor Devices, Wiley, New York (1967); S. M. Sze, Physics of Semiconductor Devices, 2[nd]ed., Wiley, New York (1981); D. H. Hill, Some Properties of Semiconductor(table), Monsanto Co., St. Peters, Mo. (1971); H. Wolf, Semiconductor, Wiley, New York (1971); W. E. Beade, J. C. C. Tsai, R. D. Plummer, Quick Reference Manual for Silicon Intergrated Circuit Technology, Wiley-Interscience, New York, 1985; F. Shimura and H. R. Huff, "VLSI Silicon Material Criteria," Chapter 15 in VLSI Handbook(ed. N. G. Einspruch, Academic Press, 1985).

[a]用于态密度计算

[b]用于迁移率计算

[*]最近有几篇文章报导硅中的本征自由载流子浓度 $n_i = 1\times10^{10}cm^{-3}$ M. A. Green, J. Appl. Phys., vol.67, p. 2944(1990); A. B. Sproul and M. A. Green, J. Appl. Phys., vol.70, p. 846(1991), and S. M. Sze, Semiconductor Devices: Physics and Technology (2[nd] ed. Wiley, New York 2002), pp. 36-37] 然而,表中给出的本征载流子浓度 $n_i = 1.45\times10^{10}cm^{-3}$,使用范围更广泛,在习题和本书的讨论中也采用这个值。

表 1.4 硅的其他性质(轻掺杂,除特别指出外,均为 300K)

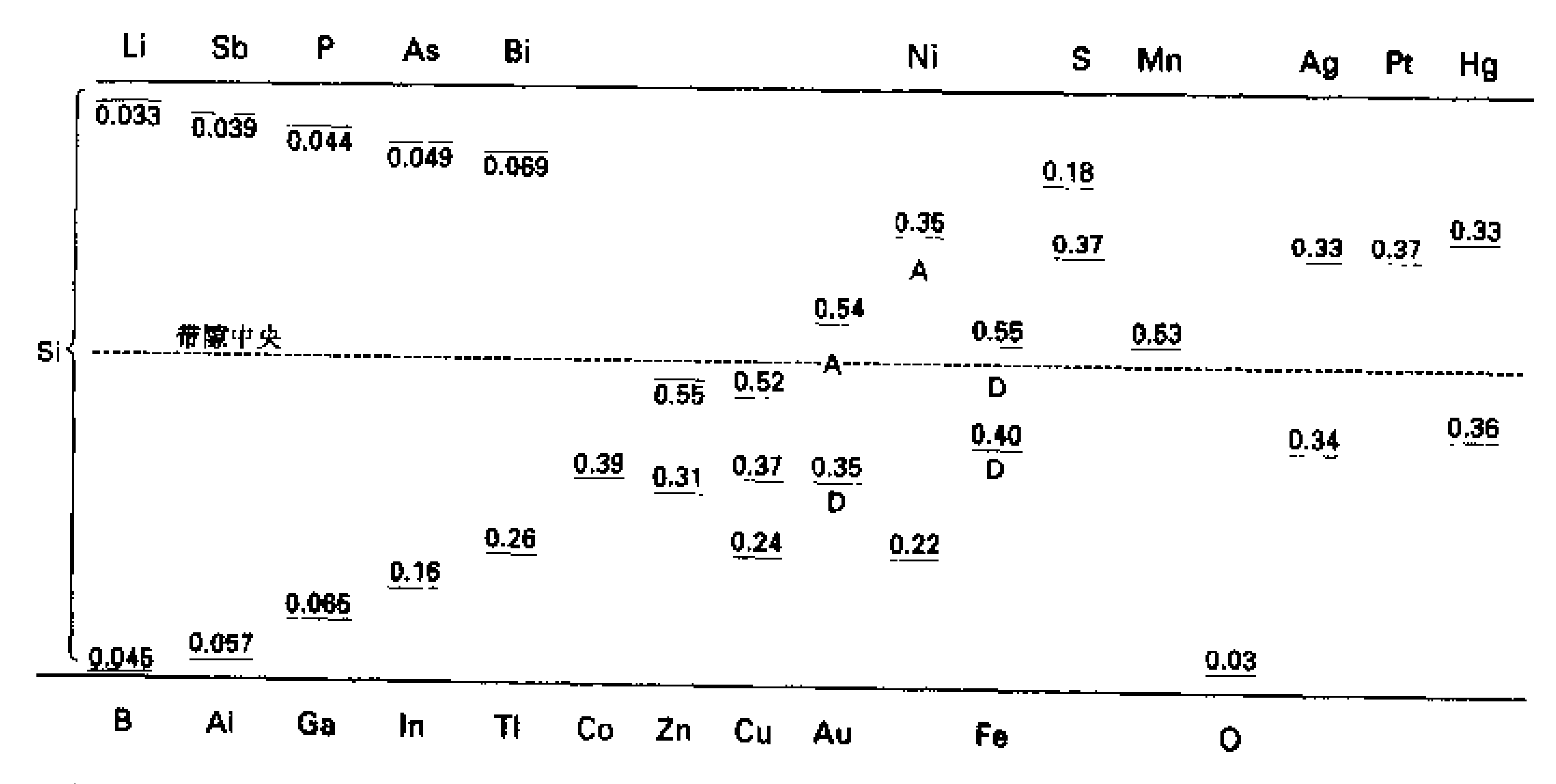
性质	符号	单位	值
四面体半径		nm	0.117
禁带宽度的压力系数	$\Delta E_g/\Delta p$	$eV(atm)^{-1}$	-1.5×10^{-6}
本征载流子浓度[a]	n_i	(T的单位为K)	$3.87\times10^{16}T^{3/2}\exp\left[\frac{-7014}{T}\right]$
禁带宽度	E_g	eV	$1.17-\frac{4.73\times10^{-4}T^2}{T+651}$
电子迁移率的温度系数	$\Delta\mu_n/\Delta T$	$cm^2(V\ s\ K)^{-1}$	-11.6
空穴迁移率的温度系数	$\Delta\mu_p/\Delta T$	$cm^2(V\ s\ K)^{-1}$	-4.3
电子扩散系数	D_n	cm^2s^{-1}	34.6
空穴扩散系数	D_p	cm^2s^{-1}	12.3
硬度	H	Mohs	7.0

续表

性 质	符 号	单 位	值
弹性模量	c_{11}	dyne cm^{-2}	1.656×10^{12}
	c_{12}		0.639×10^{12}
	c_{44}		0.796×10^{12}
杨氏模量(<111>方向)	Y	dyne cm^{-2}	1.9×10^{12}
表面张力(1412℃)	σ_0	dyne cm^{-2}	720
熔融潜热	H_f	eV	0.41
凝固时的膨胀率		%	9.0
晶格振动截止频率	ν_0	Hz	1.39×10^{13}

主要来源:H. F. Wolf, Semiconductor, Wiley, New York (1971), p. 45

硅中元素杂质的能级[b]



[a]高温下 n_i 与 T 的关系见图 2.10。

[b]除了标注 D 的施主能级外,带隙中央以下的能级从价带顶开始算起,称为受主能级。除了标注 A 的受主能级外,带隙中央以上的能级从导带底开始算起,称为施主能级[10]。

第2章 硅 工 艺

成功的工程技术建立在两个基础之上，一个基础是要掌握基本的物理概念，同样重要的第二个基础是要有完美的加工技术，即要有将工程设计的概念转变成有用结构的手段。在第1章中，我们已经回顾了集成电路电子学中要用到的物理原理，本章将讨论在硅晶体上制造器件的工艺技术。硅集成电路工艺非常之强大，以至整个现代世界都感受到了它的影响。本章我们除了介绍目前正在使用的IC工艺外，还试图介绍这一快速发展领域的前景。

集成电路工艺的发展始于对二极管电学特性的理解和20世纪40年代末期晶体管的发明。当时，锗是最受关注的半导体材料。从锗的实验中，人们获得了具有化学纯度的大尺寸单晶生长和关于晶格完整性的重要知识。

锗是金刚石结构的晶体，晶格中的每个原子与其最近邻的4个原子形成共价键。晶体结构如图1.8所示。锗的带隙为0.67eV，300K时，本征载流子浓度为$2.5\times10^{13}cm^{-3}$。由于锗的带隙相对比较小，其本征载流子浓度随温度升高而迅速增加，400K时大约增加到$10^{15}cm^{-3}$（见图1.9）。

当本征载流子浓度可与掺杂浓度相比拟时，大多数器件都不再能正常工作，所以锗器件只能工作在温度低于70℃（343K）的环境。早在1950年，锗器件工作温度的限制就激发了对具有类似晶格结构，但能在更高温度下应用的其他半导体材料的研究。随后的几十年中，集成电路工艺的发展已集中在元素半导体硅（$E_g=1.12eV$）和化合物半导体砷化镓（$E_g=1.42eV$）。绝大多数集成电路用硅制造，而化合物半导体，例如砷化镓，用于制造专用的高性能电路。例如，具有直接带隙的化合物半导体发光效率很高，在制造光学器件时尤其有用。

本章大部分内容涉及硅工艺，但是与化合物半导体有关的工艺将在本章后半部分简单地介绍。表1.3总结了锗、硅、砷化镓和其他几种有用的电子材料的性质。表1.4列出了硅的其他一些性质。

2.1 硅平面工艺

除了具有良好的半导体特性之外，硅能获得广泛应用的主要原因是在硅上面能够形成稳定、可控且绝缘性能优良的氧化物薄膜（二氧化硅SiO_2），这种能力是任何其他半导体-绝缘体系统不能相媲美的。利用氧化层可在硅样品上选定小的区域进行可控掺杂，同时氧化物又能阻止其余区域的掺杂。能够在硅上选定的小区域掺杂是制造集成电路中密集的器件阵列的关键。

Si-SiO_2系统的两个化学性质对硅工艺非常重要。第一，利用液态或气态刻蚀剂可选择性刻蚀其中一种材料。例如，氢氟酸能溶解二氧化硅却不腐蚀硅。第二，二氧化硅层能掩蔽下面的单晶硅，使其不受来自高能离子束或高温气相扩散源的杂质原子的影响。

利用这些特性，可将杂质原子掺入到硅片上没有被厚二氧化硅掩蔽的区域。经掩膜板曝光的光敏聚合物薄膜，可以准确地限定掩蔽区域。当硅片浸入氢氟酸浴槽或暴露于气相刻蚀剂中时，聚合物图形能保护硅片表面选定的氧化物区域，从而形成了二氧化硅层中有裸露硅窗口的表面。这种选择性刻蚀工艺，最初是为平板印刷而开发的，可刻画出非常小的图形。当把硅片放置到淀积杂质原子的环境中时，这些杂质原子就只能进入裸露的硅窗口中。

将刚才描述的氧化、图形化、掺杂操作进行适当的排序和重复，可以将p型和n型掺杂原子选择性地引入到小至数百纳米尺寸的区域内。这些工序是硅平面工艺的基本步骤。因为它们是通过在硅晶体表平面附近实施一系列工序来产生器件结构的，所以被称为平面工艺。

除了提供一种限定掺杂区域的手段外，在硅上形成的高质量氧化层还能提高硅衬底表面的电学性能。由于硅衬底上周期性晶格结构的中断，在理想的自由表面上有不完整的或者说悬挂键存在。我们将在第4章中看到，这些断裂键会在硅衬底表面的能带带隙中形成允许态，并且会使靠近表面的器件区的电学性能退化。但是，如果在硅表面上形成高质量的二氧化硅，就可钝化几乎全部的这些表面态，从而在硅表面区获得几乎理想的性能。虽然硅表面处的键面密度约为$10^{15}cm^{-2}$，通过在表面上适当生长一层二氧化硅，可使具有电活性的键数目减少至小于$10^{11}cm^{-2}$。能够在硅表面消除几乎所有的电活性态，从而使无处不在的硅金属-氧化物-半导体(MOS)晶体管很好地工作，这正是现在大多数大规模集成电路的基础。

用于平面工艺的硅晶体是由大尺寸单晶硅制备的薄片，称为硅(晶圆)片。通常先将掺杂原子淀积到硅片表面上或表面附近选定的区域，然后将它们扩散进入硅中。因为杂质原子从表面引入，且通常扩散尺寸非常小，所以平面工艺的器件有源区深度只在硅片表面几个微米范围内。硅片的其余厚度(通常几百微米)只用做重要的表面区的机械支撑。

平面工艺的一个主要优点是每一个制造工序(在封装之前)都是对整个硅片进行的。因此，能够将许多器件高精度地互连在一起构建成一个集成电路(IC)。现在，单个集成电路尺寸通常可达20mm宽，所以一个硅片(大部分直径为20~30cm)可以容纳许多集成电路。从经济的角度而言，显然要求不断增大每个硅片的面积并同时减小硅片上每个集成电路的面积。

平面工艺中的最重要的工序如图2.1所示，包括(a)氧化物掩蔽层的形成，(b)氧化物的选择性去除，(c)在硅片表面或其附近淀积杂质原子，(d)杂质原子扩散进入裸露硅的区域。这些工艺确定了掺杂的位置，从而决定了器件和集成电路的电学性能。

平面工艺和与之相关的硅基电子学一直在持续高速地发展。图2.2给出了过去40年硅集成电路工艺发展的部分成果。

集成电路的最小特征尺寸继续快速不断缩小，从1969年的8μm减小到当前的130nm[1]。通过绘制最小特征尺寸(用对数刻度)和首次实现商品化生产年度的关系曲线就可以发现发展的速度，如图2.2(f)所示。在该图中直线显示最小特征尺寸按指数减小。这种指数变化规律首先被Intel公司的Gordon Moore发现并以"Moore定律"著称于世。(Moore定律已被人们普遍接受并不断引用，甚至在卡通片"Dilbert"中也提到了它[1])。随着特征尺寸的迅速减小，IC芯片的尺寸在不断地增大，但是芯片增大的速度小于最小特征尺寸减小的速度。由于将更小的特征尺寸和更大的芯片相结合，使得一个IC芯片上的晶体管数目增加的速度更快——每18~24个月增加1倍(见图2.2(g))。持续的工艺发展使这一快速的增长成为可能，而晶体

[1] $1nm = 10^{-9}m = 10^{-7}cm = 10^{-3}\mu m$

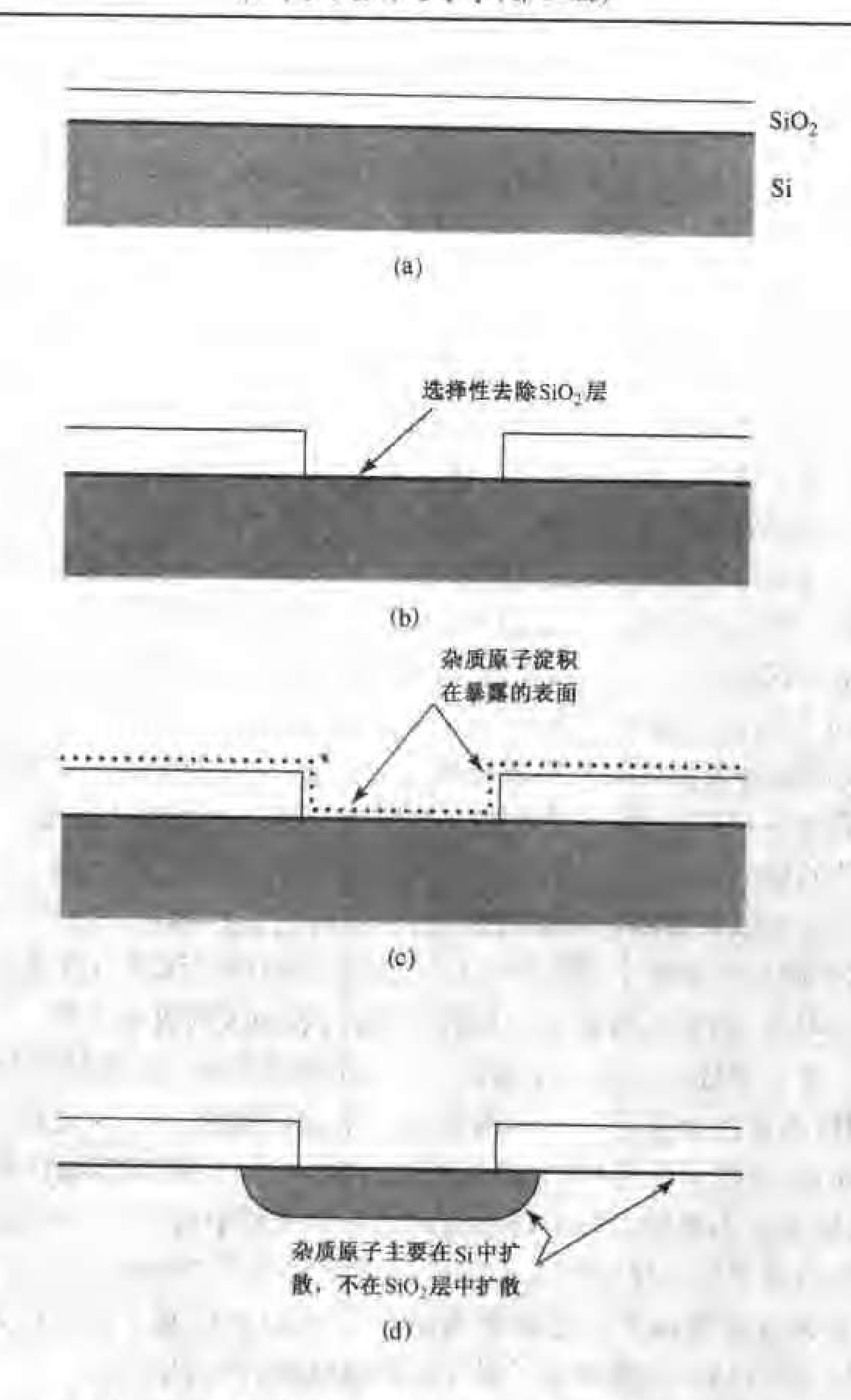
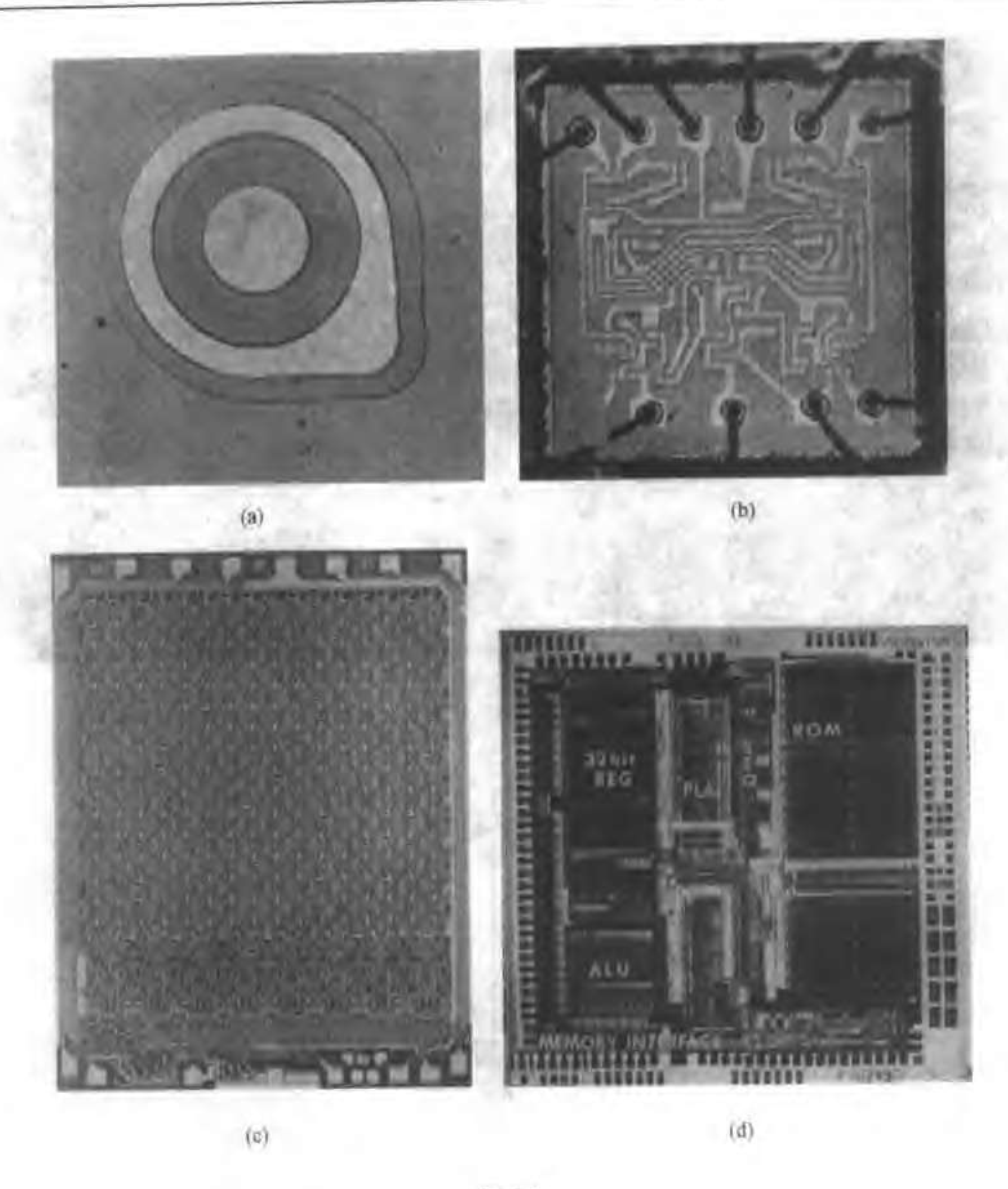


图2.1 硅平面工艺的基本制作过程。(a)二氧化硅的形成
(b)选择性去除二氧化硅；(c)在硅片上淀积杂质原子；
(d)杂质原子扩散进入裸露的硅区域

管的基本物理仍然相对保持不变(但是，在大尺寸晶体管中一些不重要的二级效应在更小晶体管中起主要作用)。

因为生产更小特征尺寸的电路需要不同的制造设备，所以特征尺寸不是连续减小的。当然，对于每一代器件的晶体管面积下降的比例要能带来足够的利益以弥补购买新设备所需的开销。通常是将面积减小一半、线条尺寸减小1/4为原来的0.7倍。现在60nm器件已经生产，20~30nm器件也已被验证。

Moore定律已被大家普遍接受，它几乎成为一个总会实现的预言。对图2.2(f)中过去特征尺寸的发展规律进行外推，可以确定未来每年最小特征尺寸的发展目标。对特征尺寸和其他物理及电学性能的预测在国际半导体技术发展蓝图(ITRS)[3]中已被指出，并且在不断更

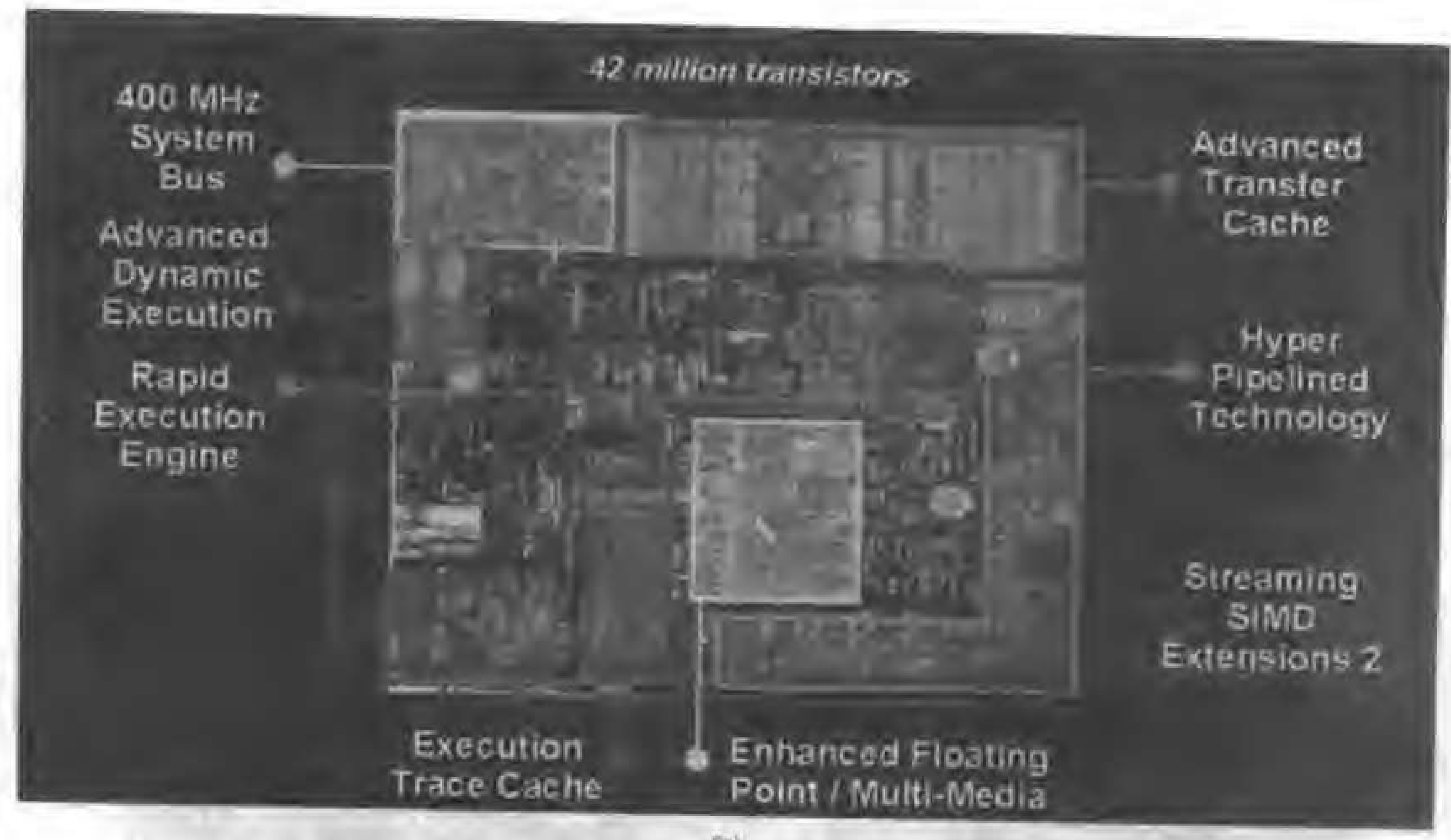


(a) (b) (c) (d)

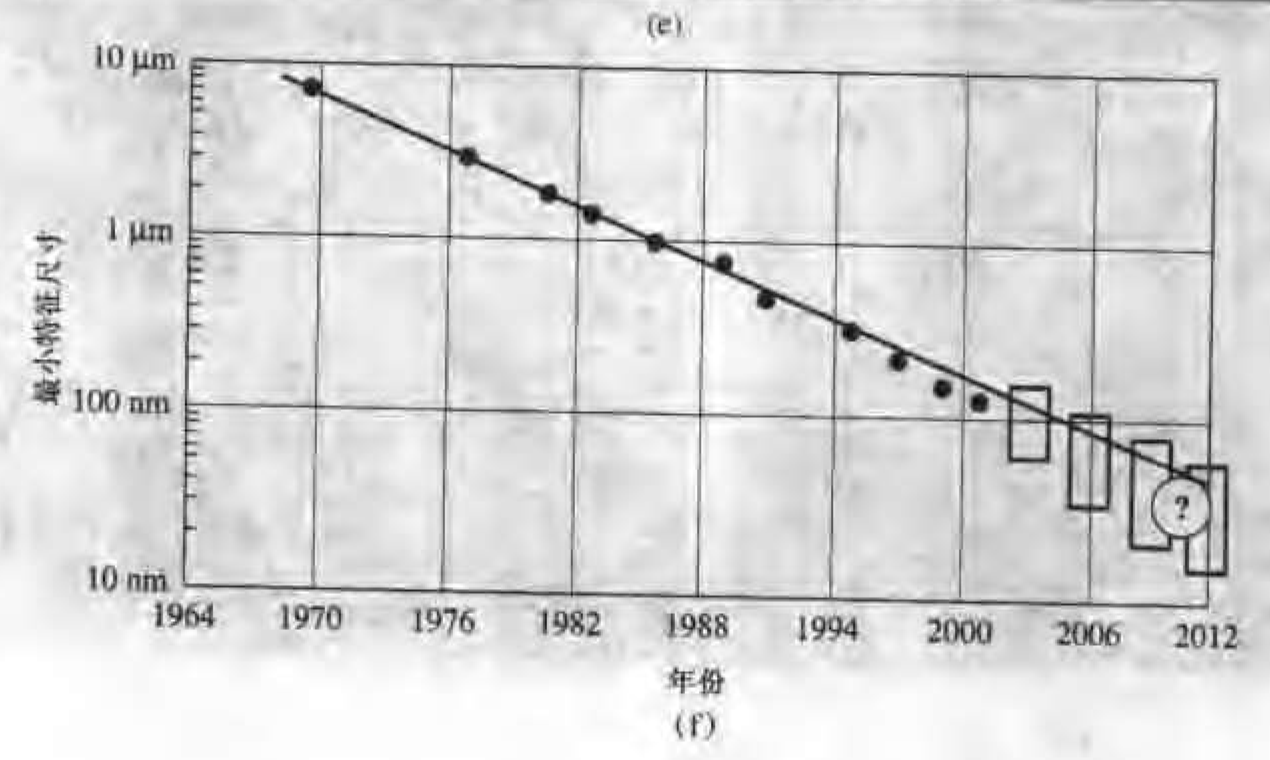
图 2.2

新。半导体制造商将投资研发预定的下一代最小尺寸特征的技术。实际上，变化有时会超过预测的发展速度。如果预定尺寸对工业界是一个标准，那么每个公司都试图研发比预测更快的技术以便在市场中更具有竞争力。

但是器件不可能无限制地按比例缩小下去。它将受到两个因素的制约。第一，随着器件特征尺寸不断下降，每个晶体管中的电子数目进一步减少，如图 2.2(h)所示。随着电子数 n 的降低，统计涨落值(约为$\sqrt{n}$)在总电子数中的比例会增加，这将限制电路的性能并使电路更难以进行设计。早在几年前，这些统计涨落已经开始影响模拟电路的性能。在几年内，类似的统计波动将会影响到数字电路的设计。随着统计波动重要性的不断增加，我们终将达到每个晶体管中只有一个电子的时代(可能在 2015 年)。那时，电子器件的整个概念不得不发生变化。在一些先进的实验室中正在研究可能替代晶体管的其他结构，但至今还没有找到。



(e)



(f)

图 2.2 续

第二,即使考虑眼前的常规器件,每一代工艺的发展都变得更困难且更昂贵。用于定义最小特征尺寸的光刻技术通常是限制因素。最终,技术研发和制造新设备的成本可能会限制 IC 技术的进一步发展。目前一套现代集成电路设施价值几十亿美元,而且成本还在不断增长。制造设备的高成本限制了承担制造集成电路任务的公司数量。事实上,许多公司设计的电路是由代工厂制造的,这些代工厂专为其他公司大批量制造集成电路。

但是,即便存在这些限制,平面工艺将继续在电子学领域占主导地位若干年。因为它为计算机、通信和日用消费品带来了巨大的利益,导致有巨额资金不断投入在研究、开发和制造设施方面。平面工艺是硅集成电路生产的基础,也是保证不断发展的日益复杂的电路能够进行生产的基础。要能够较好地理解和设计器件,就必须对硅工艺的基本过程有相当全面的了解。本章中的许多内容直接提供了这方面的知识,也是随后几章讨论器件的基础。

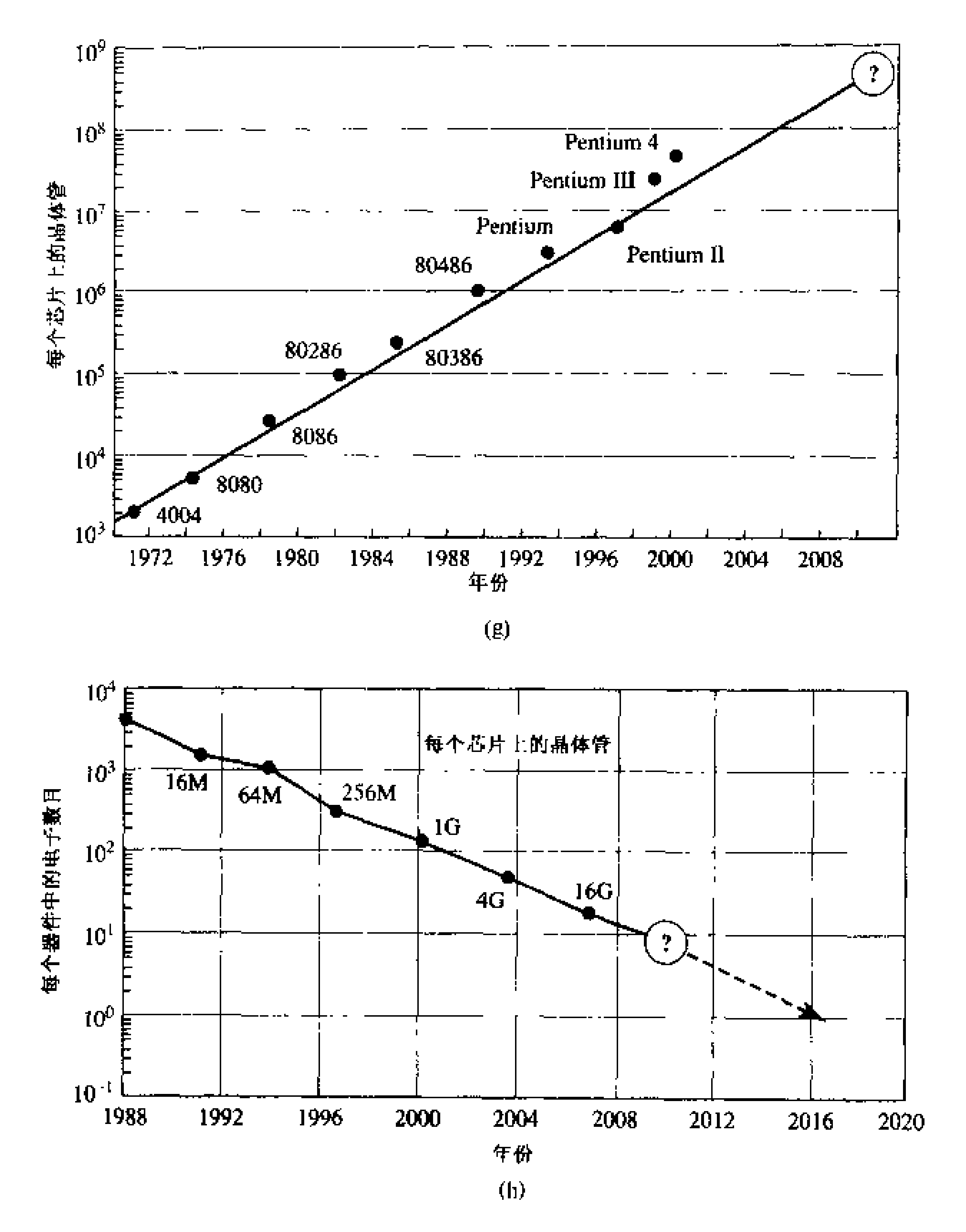


图2.2 IC工艺的发展。(a)第一个商用的硅平面晶体管(1959)(外径为0.87mm);(b)二极管-晶体管逻辑电路(DTL)(1964)(芯片尺寸为1.9mm);(c)256位双极随机存储器(RAM)电路(1970)(芯片尺寸2.8mm×3.6mm);(d)含有450000个晶体管的VLSI计算机中央处理器芯片(1981),图中标出了不同功能的IC模块(芯片尺寸为6.3mm²)[(a),(b),(c)引自B. E. Deal - Fairchild Semiconductor,(d)引自Hewlett-Packard Co.];(e)含有4千2百万晶体管的奔腾4处理器(2000)的结构图;相应的芯片照片见本书的封面(引自Intel Corporation);(f)最小特征尺寸与第一个商业产品出现的年度之间的关系;(g)Moore定律的另一种表达形式,约30年来芯片上的晶体管数目每18~24个月翻一番;(h)随着特征尺寸的减小,每个器件中的电子数目下降[(f)~(h)的数据来自Mark Bohr, Intel;Howard Huff, Sematech;Joel Birnbaum, Hewlett - Packard; Motorola]

2.2 晶体生长

制造集成电路要使用大尺寸高质量的单晶硅。下面简单考虑一下对硅的基本要求有助于更好地理解“高质量”的含义。

在典型的集成电路应用中,要求掺杂浓度大约在 $10^{15} \sim 10^{20} cm^{-3}$ 范围内。为了控制器件性能,任何非有意掺入的或本底电活性杂质浓度应低于最小有意掺杂浓度至少2个数量级,这就是说约为 $10^{13} cm^{-3}$ 或更低。因为硅的原子浓度为 $5 \times 10^{22} cm^{-3}$,也就是说,每十亿个硅原子中含有一个非有意掺入的电活性杂质是可以容忍的,这么高的纯度远远超过其他工业对原材料的要求。制造集成电路需要这么高质量的材料证明了硅工艺是极其精细的。

高纯硅是由两种普通材料制成的:二氧化硅(存在于普通沙子中)和元素碳。在高温(约2000℃)电弧炉中,碳将二氧化硅还原成元素硅,冷凝后成为纯度约为90%的冶金级的硅,这时的纯度还不能用于制造半导体器件。将冶金级硅转变成液态的三氯氢硅($SiHCl_3$)可实现进一步提纯。选择性蒸馏法(分馏法)将三氯氢硅与其他氯化合物分离开来。然后纯化后的三氯氢硅被氢还原成高纯度半导体级的固态硅。这时的硅是多晶体,由许多随机取向的小晶体组成。这种元素态多晶硅通常被淀积在高纯半导体级硅棒上以避免被污染。

由于晶粒边界和其他晶体缺陷会降低器件的性能,因此最后需要制备出大尺寸(直径约为20~30cm)且几乎完美的单晶硅,要获得这样高质量的单晶硅需要先进的设备。利用直拉法和区熔法可以制备出这样的晶体。

直拉法(Czochralski 法) 制备集成电路衬底材料最常采用的是直拉法。这种方法是在稍高于硅的熔点温度(1412℃)和惰性气氛(通常是氩气)下,首先把多晶硅棒放在石英坩埚中熔化,如图2.3所示。

然后,将具有所需晶向的高质量籽晶一边旋转一边降落到熔融硅中,如图2.3(a)所示。同时,坩埚向相反方向旋转以混合熔融硅并将温度的不均匀性降至最小。一部分籽晶在熔融硅中溶解以去除有应变的外层而露出新鲜的晶体表面。接下来慢慢地将籽晶从熔融硅中升起(或提拉)。随着籽晶的升高和冷却,来自熔融硅的材料在籽晶上凝结形成更大的晶体,如图2.3(b)所示。在生长期间,需要一直精确控制生长条件,这样新的硅原子会继续按照已经凝固的晶体结构结晶。利用自动反馈机构控制提拉速度和温度就可以获得需要直径的晶体。以这种方式能制造出圆柱体单晶硅。随着晶体生长技术的发展,圆柱形晶体的直径已经从几毫米不断增大到现在的20cm或30cm。

对于许多集成电路工艺而言,要求硅的起始掺杂浓度约为 $10^{15} cm^{-3}$。通过控制掺入熔融硅的硼或磷等所需杂质的量,就能获得这样的掺杂浓度。通常每公斤硅中必须加入重约0.1毫克的掺杂元素。为了便于精确控制,通常是将少量重掺杂硅而不是元素态杂质加入未掺杂的熔融硅中。由于硅结晶时杂质原子会受到排斥,拉制的单晶硅总是比熔融硅的掺杂浓度小。这种分凝现象导致熔融硅中的掺杂浓度随着晶体生长而不断增加,单晶硅锭中籽晶一端比尾部一端的掺杂浓度小。另外,在直拉法生长的硅晶锭中,沿径向方向也会存在微小的掺杂浓度梯度。

由于容纳熔融硅的石英(二氧化硅)坩埚会缓慢熔解,直拉硅中会含有一定量的氧。对制造集成电路的具有中等掺杂浓度的硅片来说,氧对净掺杂浓度并没有很大的影响。直拉法生

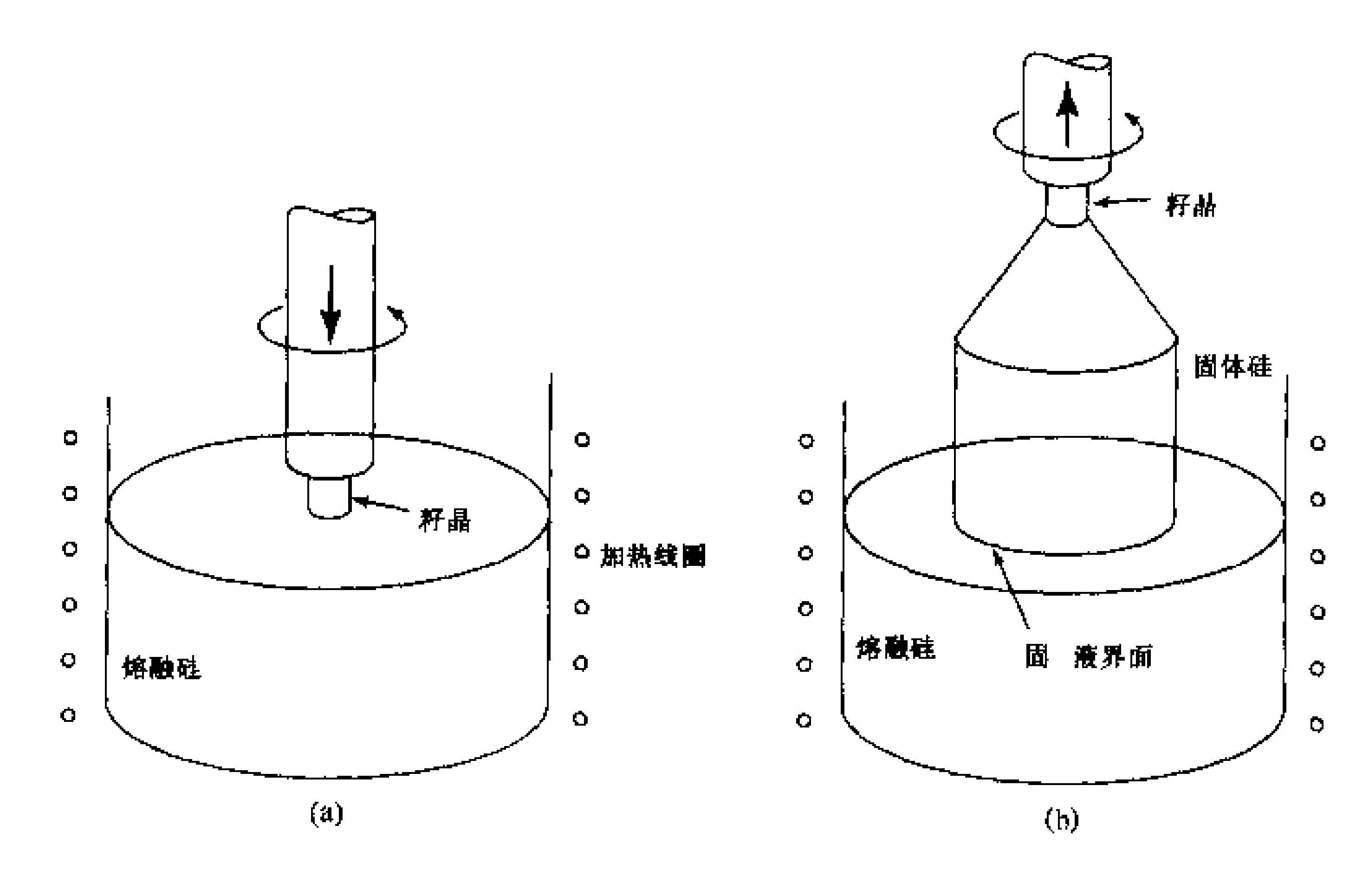


图2.3 直拉工艺制备单晶硅半导体晶锭。(a)籽晶位于熔融硅表面开始晶体生长;(b)提拉籽晶拉出单晶硅

产的单晶硅中通常包含浓度为$10^{18}cm^{-3}$的氧原子(结晶温度下氧在硅中的固溶度),它们能被用来控制集成电路硅片中非有意掺杂的杂质(通常是金属)的运动。在集成电路制造过程中,氧可能会在其他杂质趋于积聚的位置析出。如果析出的氧位于有源器件内部,就会降低器件的性能。但是,如果远离有源器件,析出的氧能吸引不需要的杂质离开有源区,起到吸杂剂的作用,从而提高器件的性能。要使高密度集成电路中的器件性能一致性好,控制氧析出的位置和尺寸是很重要的。

熔融硅的对流会加快对坩埚的腐蚀,增加氧在熔融硅和结晶硅中的氧含量。对流也会影响晶体生长的稳定性,降低晶体的质量。磁场能够抑制熔融硅中的对流。这种磁限制的直拉生长技术为改善晶体生长工艺和提高制备出的晶体的纯度提供了可能性。

硅片中的一小部分氧,约占总数的0.01%,在中等温度下退火后能起到施主的作用。这么小的施主浓度(大约$10^{14}cm^{-3}$)不会影响大部分集成电路中硅的电阻率。但是对制造功率器件和其他特种用途的高电阻率(20~100Ω·cm)硅而言,氧的问题不能忽略。因此,高电阻率硅通常用区熔工艺制造。

区熔法 在区熔工艺中,一根多晶硅柱竖直放置并旋转,而熔区(1~2cm)缓慢地从多晶硅柱的底端移动到顶端,如图2.4所示。熔区被加热(通常用射频(RF)感应加热器),并从多晶硅柱籽晶一端开始结晶,逐渐移到另一端。杂质在熔融区分凝,从而使结晶硅纯化。与直拉法制备的硅相比较,由于没有将硅放到含氧的坩埚中,所以在硅熔区中没有引入氧。虽然区熔工艺生产成本更高,且只限于制备直径较小的硅晶锭,但是硅中含氧量只有直拉法的1%,并且其他杂质浓度也降低了。多次进行区熔再结晶能制备电阻率高于$10^4Ω·cm$的单晶硅锭。

硅片制作 当生长出单晶硅锭并研磨至精确的直径后,用金刚石锯可将其切割成薄的圆形硅片。对硅片要进行化学腐蚀去除切片时造成的损伤,然后不断用更细的抛光粉和化学腐蚀剂进行抛光,直至获得消除了缺陷像镜子一样平整光亮的表面。这时的硅片就可用于器件

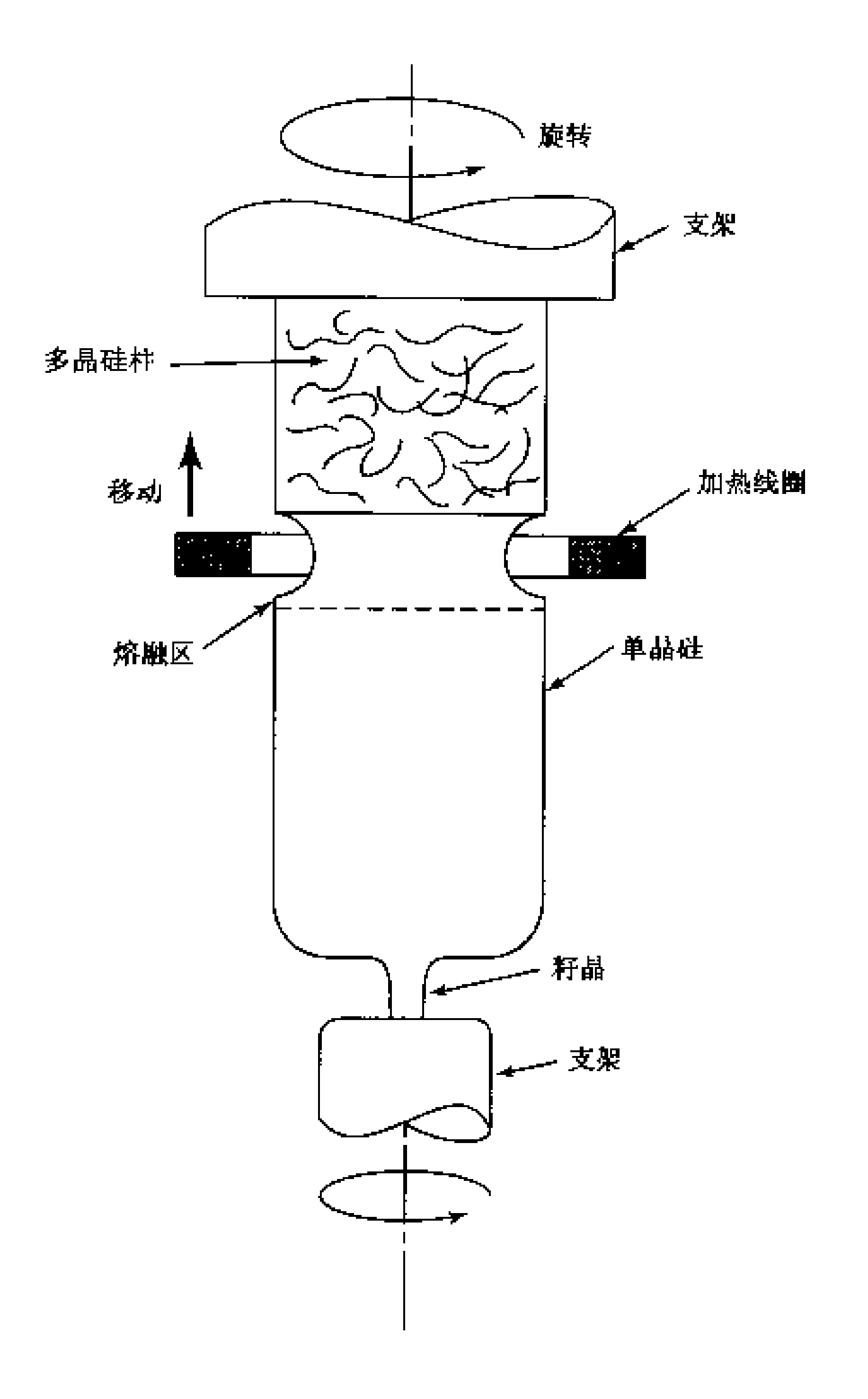


图 2.4 区熔工艺。熔融区域通过多晶硅柱，单晶硅从底端的籽晶开始生长

制造。

在切片之前，要在硅片上制作定向标记，以便沿指定晶向制造电路。特别是硅片上构成集成电路单元或芯片的四个边通常要沿着硅片容易解理的方向，以便完成平面工艺后能将芯片相互分离。分离过程通常是用锋利的针尖在芯片之间划线从而使它们分开，所以对易于解理的平面的定向是很重要的。对于比较小的硅片，采用在硅片上研磨出一个平边的方法来指示晶向，这个平边通常垂直于一个容易解理的方向。在大多数情况下，沿⟨110⟩方向形成主平边。有时要增加一个较小的副平边以确定硅片的导电类型和晶向，如表 2.1 所示。对于较大的硅片，磨制出硅片平边会明显减小硅片面积和硅片上芯片的数量。因此，在大硅片上省去了平边，而是在硅片的边缘刻一个凹口用于指示晶轴方向。

表 2.1 硅片副平边的位置

晶向	导电类型	副平边(相对于主平边)
(100)	n	180°
(100)	p	90°
(111)	n	45°
(111)	p	没有副平边

切片后，通常用激光将硅片表面一小块面积上的硅蒸发掉以制作该硅片的标记号码。这个标记指明了制造商，硅锭掺杂类型和晶向。其他数字对于各硅片是惟一的，可在硅片制造工艺中的任何一个工序中识别硅片。激光标记既能被设备操作员识别，也能被自动化的工艺设备识别。由于激光标记是在硅片被腐蚀和抛光前制作的，所以标记过程中产生的应变可在化学腐蚀时被去除，抛光溅洒出的材料也不会污染晶面表面。

2.3 热氧化

在室温下的空气中，裸硅表面会很快形成厚约2nm的氧化层。在选择性掺杂过程中，需要采用热氧化或淀积的方法制备更厚(8nm至1μm)的二氧化硅来保护硅表面。当用淀积方法形成二氧化硅时，硅和氧被输运到硅片表面并在那里反应(见2.6节)。在热氧化过程中，硅片表面的硅原子与高温炉中的氧直接发生反应。热生长法一般比淀积法制备的氧化物质量高。尽管这个氧化物属于非晶结构，但通常具有精确的化学剂量比(SiO_2)，并且与硅表面具有强的键合力。硅与热生长SiO_2之间的界面具有稳定可控的电学性能。我们将在第8章看到，具有优异品质的半导体-绝缘体界面是成功制造金属-氧化物-半导体(MOS)晶体管的基础。

热生长氧化物，首先要将硅片放在石英管内，这个石英管固定于电阻加热炉圆柱形孔中。炉子可以水平安放，如图2.5所示，也可以垂直安放。硅片表面一般垂直于主要气流方向。典型温度是850℃～1000℃，温度越高反应进行的速度越快。虽然当温度达到1412℃时硅才会熔化，但是为了减少晶体缺陷的产生和前面已引入的杂质原子的移动，实际的氧化温度要低得多。另外，超过1150℃后，石英炉管和其他固定装置开始软化并且性能退化。

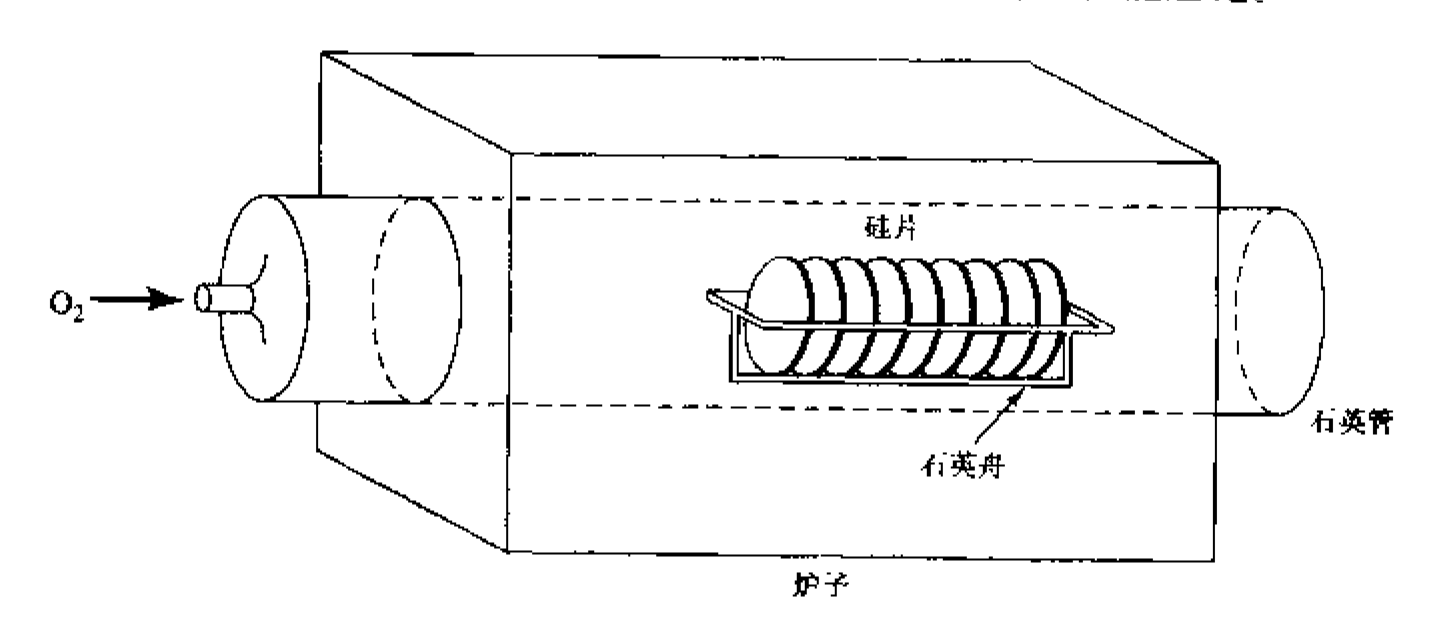


图2.5 在高温炉中硅片与氧化气氛接触生长二氧化硅绝缘层

氧化气氛可以是干氧，或者是包含水蒸气的气氛，在高温炉中一般采用氧气与氢气反应来产生水蒸气。使用这种高温合成的蒸汽需要慎重的安全程序来处理易爆的氢气，通常引入少许过量的氧以避免在炉中留下未反应的氢气。将高纯度的干氧或氮气从加热到接近沸点的水中穿过也能形成蒸汽环境。整个氧化反应过程为

$$Si(s) + O_2(g) \rightarrow SiO_2(s) \qquad (2.3.1a)$$

和

$$Si(s) + 2H_2O(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{2.3.1b}$$

在蒸汽气氛中氧化的速度要快得多,因此一般用于制备厚的二氧化硅保护层。在干氧气氛中生长厚的二氧化硅,由于生长速度较慢,会使前面工序中掺入硅晶圆片的杂质产生不必要的移动(再分布)。

氧化发生在 $Si\text{-}SiO_2$ 界面处,因此氧化剂必须扩散穿过已经生成的氧化层,然后才能与界面处的硅发生反应(图2.6)。当温度较低且氧化层较薄时,$Si\text{-}SiO_2$ 界面处的表面反应速率限制了生长速率,并且氧化层厚度随氧化时间线性增长。

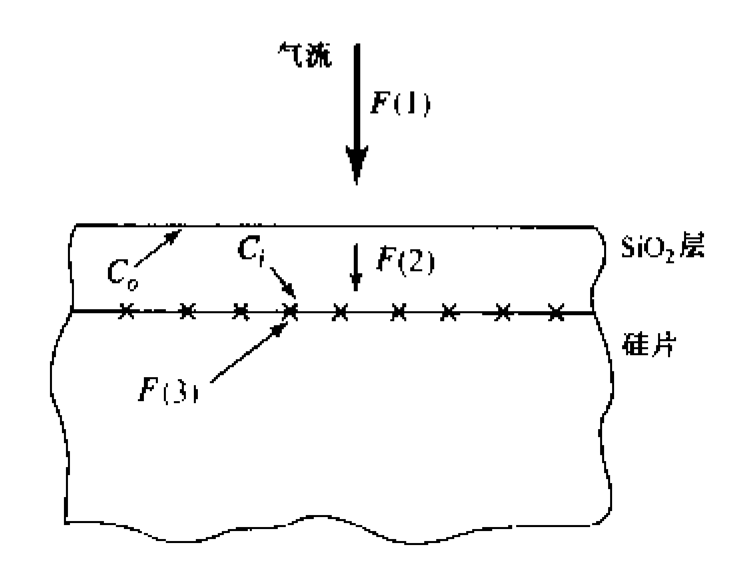


图2.6 表征氧化速率的三种流量:$F(1)$ 为到达表面的氧化剂流量,$F(2)$ 为通过已形成氧化层的氧化剂扩散流,$F(3)$ 为界面的反应流。氧化剂的浓度从氧化层中靠近气体界面的 C_o 变为靠近硅界面的 C_i

当温度较高且氧化层较厚时,氧化过程受氧化剂穿过已形成氧化层的扩散速率的限制。此时,生长的氧化层厚度大约与氧化时间的平方根成正比,这种平方根关系是典型的扩散过程的特征。

氧化动力学

对大部分氧化过程(形成遵守不同的氧化动力学的薄氧化层后),氧化层厚度 x_{ox},氧化时间 t 及温度 T 的关系可通过使以下的速率相等得到:(1)气相氧原子到达生长的氧化层的速率,(2)氧原子通过已生长的氧化层的速率,(3)根据等式(2.3.1),氧原子在 $Si\text{-}SiO_2$ 界面(图2.6)的反应速率。这些考虑构成了的经典 Deal-Grove 模型的基础[5],下面来描述。

首先,考虑气相氧化剂(氧或水蒸气)向已生长氧化层的外表面的转移。这一转移速率正比于固体表面处氧化剂的实际浓度 C_o 与 C^* 之差,其中,C^* 是气相平衡时的氧化剂浓度

$$F(1) = h(C^* - C_o) \tag{2.3.2}$$

式中,h 为气相质量转移系数,浓度 C^* 和 C_o 用理想气体定律 $C = p/kT$,与相应的气相分压 p 联系起来。

氧化剂通过已生长的二氧化硅扩散到 $Si\text{-}SiO_2$ 界面,这一过程与1.2节讨论的电子和空穴的扩散运动相似(等式(1.2.15))。氧化剂的扩散流可写为通过氧化层的浓度梯度 $(C_0 - C_i)/x_{ox}$(图2.6)与扩散系数 D 的乘积,扩散系数描述的是氧化剂在形成的氧化硅中扩

散的难易程度。C_i是氧化剂在氧化硅中靠近Si-SiO_2界面处的浓度。因此扩散流量为

$$F(2) = D\frac{(C_o - C_i)}{x_{ox}} \qquad (2.3.3)$$

氧化剂在Si-SiO_2界面处的反应用一个比例常数k_s来表征，所以氧化剂的反应速率为

$$F(3) = k_s C_i \qquad (2.3.4)$$

在稳态，$F(1) = F(2) = F(3) = F$。氧化速率可从流量中求出，用单位体积氧化物中氧化剂分子的浓度N_{ox}表示。从等式(2.3.2)～(2.3.4)中消去C_o和C_i，可得到氧化生长速率R为

$$R = \frac{dx_{ox}}{dt} = \frac{F}{N_{ox}} = \frac{k_s C^* / N_{ox}}{(1 + k_s/h + k_s x_{ox}/D)} \qquad (2.3.5)$$

解等式(2.3.5)，可得在时间t内生长的氧化层的厚度为(习题2.6)[5]

$$x_{ox} = \frac{A}{2}\left[\sqrt{1 + \frac{(t+\tau)}{A^2/4B}} - 1\right] \qquad (2.3.6)$$

其中

$$A = 2D\left[\frac{1}{k_s} + \frac{1}{h}\right] \qquad (2.3.7)$$

和

$$B = \frac{2DC^*}{N_{ox}} \qquad (2.3.8)$$

参数τ主要与初始表面氧化物的厚度有关。当氧化时间较短时，表面反应速率$F(3)$限制了氧化物的生长，等式(2.3.6)中x_{ox}与t可近似表示为线性关系

$$x_{ox} = \frac{B}{A}(t + \tau) \qquad (2.3.9)$$

等式(2.3.9)中的比例因子B/A称做线性速率系数，与界面处键的断裂有关[$F(3)$]，因此取决于晶向。IC制造中最常用的是(100)和(111)晶向的硅片。(111)晶向的线性速率系数要大一些，因为相邻的两个晶面间的键密度要低一些。

长时间氧化时，根据等式(2.3.6)，x_{ox}与t变为平方根关系

$$x_{ox} = \sqrt{B(t+\tau)} \approx \sqrt{Bt} \qquad (2.3.10)$$

等式(2.3.10)中系数B称为抛物线速率系数，与在已形成的氧化物中的扩散速率有关[流量$F(2)$]，而与硅的晶向无关。线性速率系数B/A和抛物线速率系数B的实验值如图2.7(a)和2.7(b)所示。

实际上，A、B和τ的值是从不同温度下氧化层厚度随时间变化的实验值确定的。氧化物厚度与氧化时间的关系示于图2.8(a)(干氧)和2.8(b)(湿氧)。

两个图中的数据适用于(111)和(100)硅片的氧化生长。(100)晶向的氧化层厚度要小一些，尤其在氧化温度较低和氧化层较薄时，生长受到表面氧化速率[$F(3)$]的影响。作为图2.8

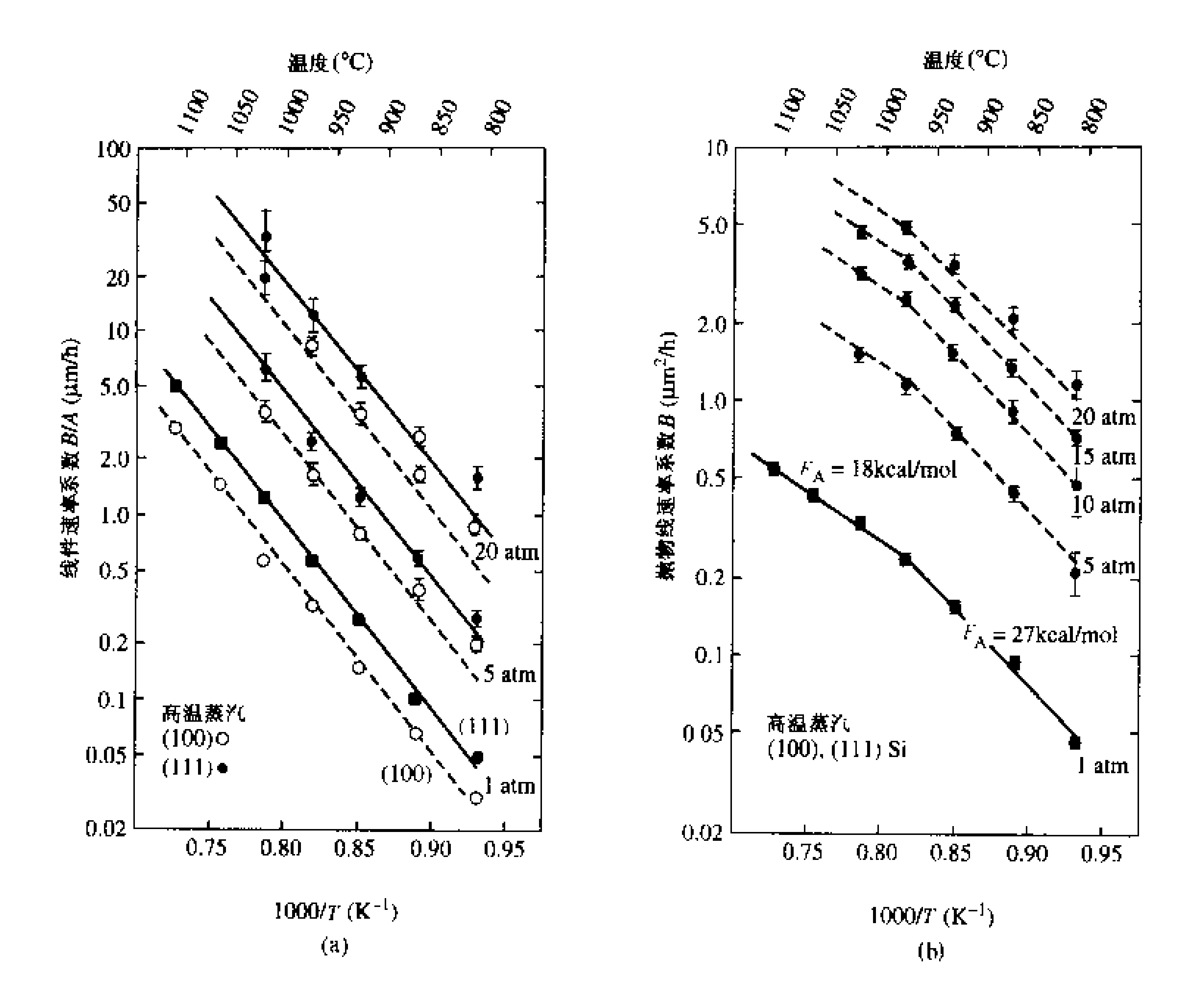


图2.7 (a)(100)和(111)晶向的硅片在1,5,20个大气压的热蒸汽中的线性速率系数B/A与$1000/T$的关系 (b)(100)和(111)晶向的硅片在1,5,20个大气压的热蒸汽中的抛物线速率系数B与$1000/T$的关系(1cal = 4.1868J)

(a)和(b)的一个应用实例,我们可以比较1100℃时在(100)硅片上生长厚度为300nm(0.3μm)氧化层所需的时间,根据图2.8(a),干氧氧化需4.4小时,根据图2.8(b),湿氧氧化需17分钟。采用常规的氧化工艺,一般的氧化层厚度为几百纳米,最多为1~2μm。

在IC制造中,经常需要确定已知厚度的氧化层继续氧化后的厚度变化关系。为此,我们可以利用等式(2.3.6)中的参数作为在继续氧化条件下生长现有氧化层厚度所对应的时间最后的厚度由加上新的氧化时间t来决定,计算得到的总厚度对应的总时间为$(t+\tau)$。

例题 计算氧化层厚度

(111)硅片上覆盖有100nm的二氧化硅,计算在1000℃干氧中氧化2小时后总的氧化层厚度。

解:由图2.8(a)知,

$$x_i = 100\text{ nm} \qquad \tau = 120\text{ min}$$
$$(t+\tau) = 240\text{ min} \qquad x_{ox} = 153\text{ nm}$$

注意氧化层厚度并不随时间线性增加。初始的100nm SiO_2在1000℃干氧中生长需2小时,而随后的2小时只生长了53nm的二氧化硅。

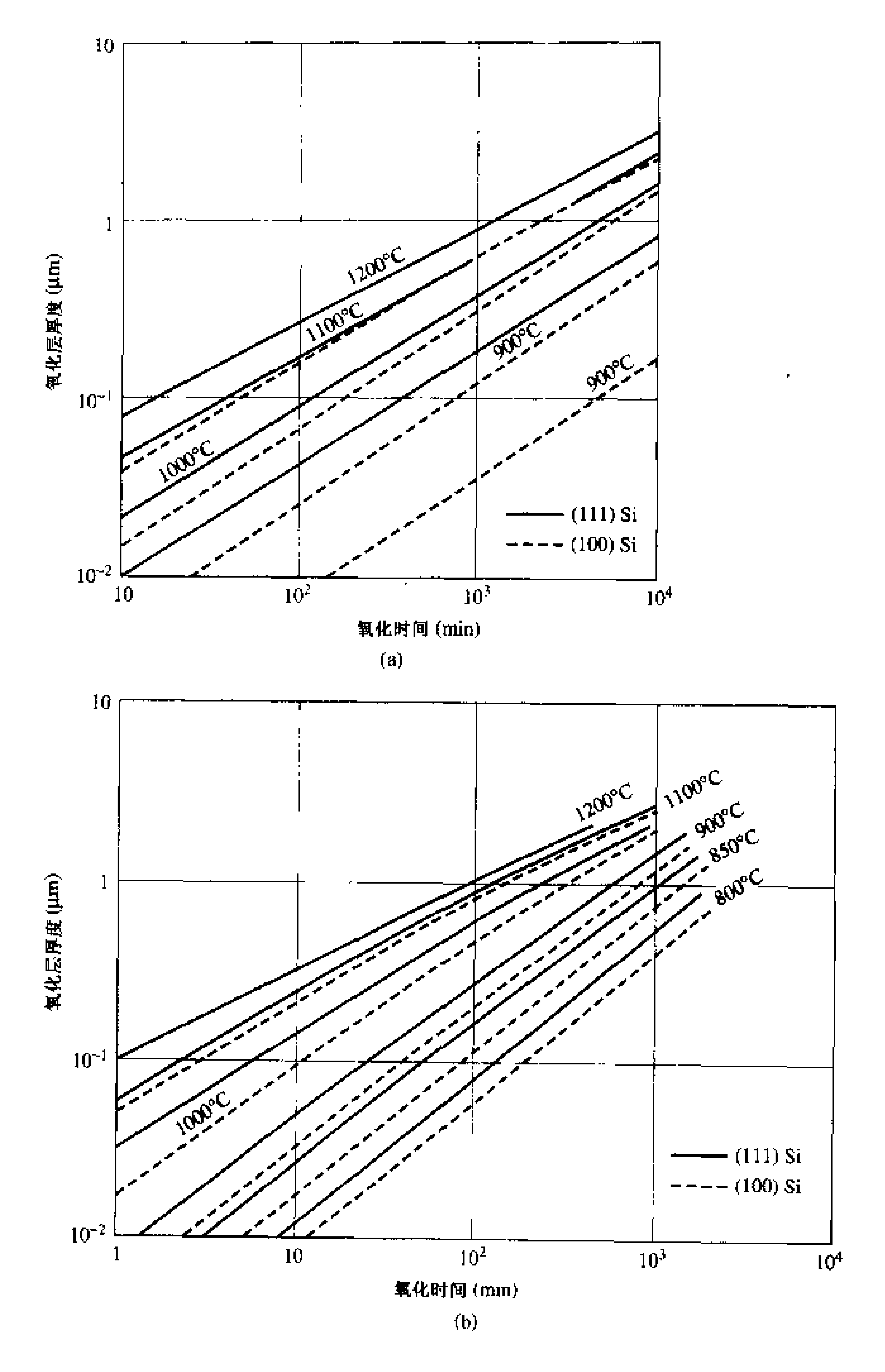


图 2.8 (a)几个常用氧化温度下，(100)和(111)硅片的氧化层厚度与干氧氧化时间之间的关系。(b)(100)和(111)硅片的氧化层厚度 x_{ox} 与湿氧氧化时间的关系(通常由氢和氧在氧化炉入口处反应得到水蒸气)。摘自文献[7]

尽管等式(2.3.5)描述了氧化速率与很宽范围内的氧化层厚度的关系，初始阶段的氧化速率要比等式(2.3.5)计算结果快得多，这一快速生长的薄氧化层可在等式(2.3.5)中加一

项来描述[8]

$$\frac{dx_{ox}}{dt} = K\exp\left(-\frac{x_{ox}}{L}\right) \quad (2.3.11)$$

其中,L 为特征长度,约为7nm的量级,K 与温度有关[8]。初期的快速氧化可能与靠近硅表面包含有其他的氧化位置有关。对厚的二氧化硅,初始的快速生长可用等式(2.3.6)中实验确定的参量 τ 来描述。

在热氧化过程中,硅被消耗形成二氧化硅。因为 SiO_2 的分子密度为 $2.2\times10^{22}cm^{-3}$(其中的硅原子浓度也等于这个数目),而纯硅的原子密度为 $5.0\times10^{22}cm^{-3}$,因此消耗掉的硅的厚度是形成的 SiO_2 厚度的0.44倍。这一关系对任意晶向的硅都成立,对多晶硅也成立,因为它只取决于体密度。

高压氧化 生长厚的二氧化硅很费时间,因为氧化层的厚度随时间的平方根增加(等式(2.3.10))。在氧化过程中由于掺杂原子会发生运动,所以通常期望快速氧化。快速氧化可采用高压氧化剂来实现。

如果 A 和 B 值已知,大气压下的氧化工艺可用等式(2.3.6)来描述,同样的方程式在其他压力下也适用(A 和 B 值不同)。厚二氧化硅的氧化速率(等式(2.3.10))由参数 B 确定。从等式(2.3.8)看出,B 随 C^* 的增加而增加,C^* 为气相平衡时的氧化剂浓度。氧化剂处于高压下对应的 C^* 可增加。典型的高气压为10至20个大气压,引起参数 B 增加10至20倍。尽管氧化剂在二氧化硅中的扩散系数 D 与压力有点关系,但对氧化速率起主要作用的是浓度 C^*。

在同样温度下,高压氧化比大气压下达到预定的二氧化硅厚度所需的时间要短得多;另一方面,如果二氧化硅厚度确定,也可在低温下以相同时间获得。采用低的氧化温度下可使热氧化过程中的晶体缺陷减少。

杂质增强氧化[†] 重掺杂n型硅比轻掺杂硅氧化速度快得多(图2.9)。IC中重掺杂和轻

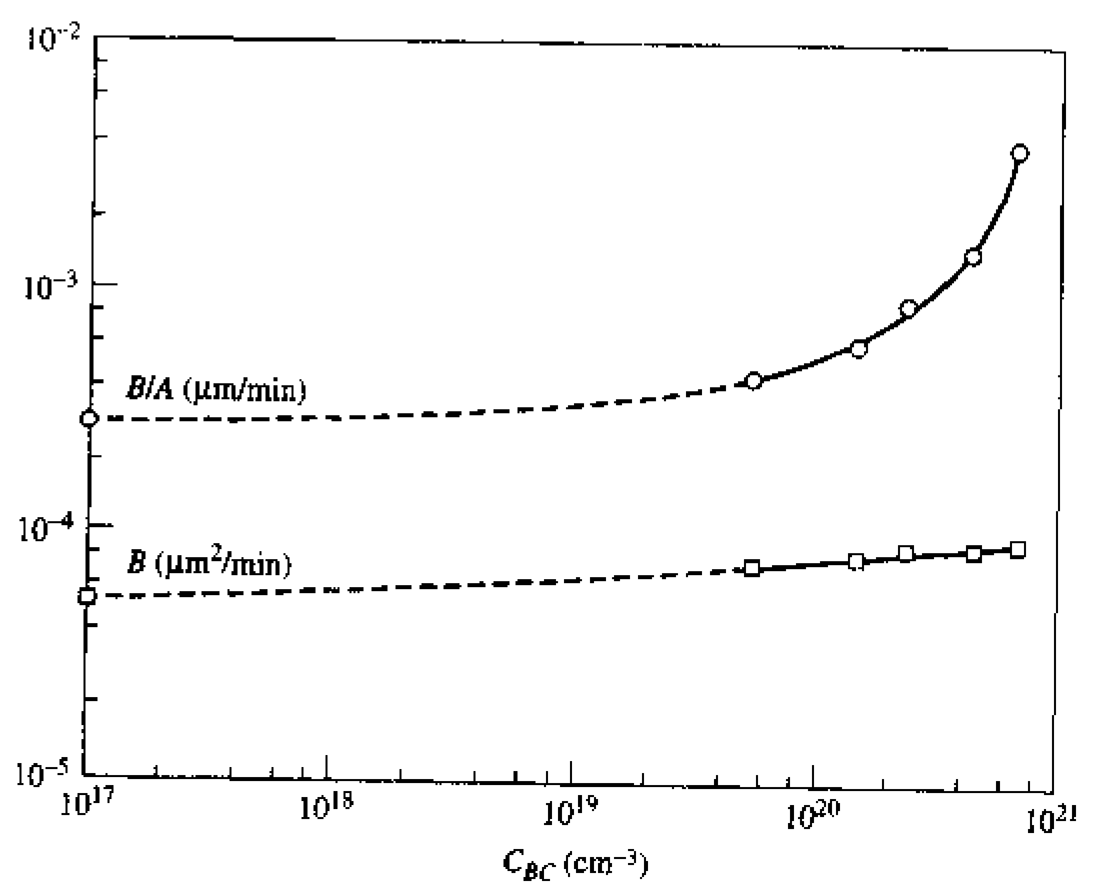


图2.9 900℃氧化时,线性速率系数(B/A)和抛物线速率系数(B)与衬底初始磷浓度的关系[9]

掺杂区不同的氧化速率可用来选择性确定想得到的区域；相反，不同的氧化层厚度又使得对电路中各部分氧化层均匀刻蚀变得困难

高浓度的某些杂质会引起硅晶格中孤立的点缺陷，它由失去的硅原子（空位）或者出现多余的硅原子（间隙原子）组成。在氧化温度下，如果这些点缺陷的浓度大于本征载流子浓度 n_i 就会影响表面反应速率（因而也影响线性速率系数 B/A）。如图 2.10 所示，n_i 随温度增加很快，在 1000℃时约为 $10^{19}cm^{-3}$。

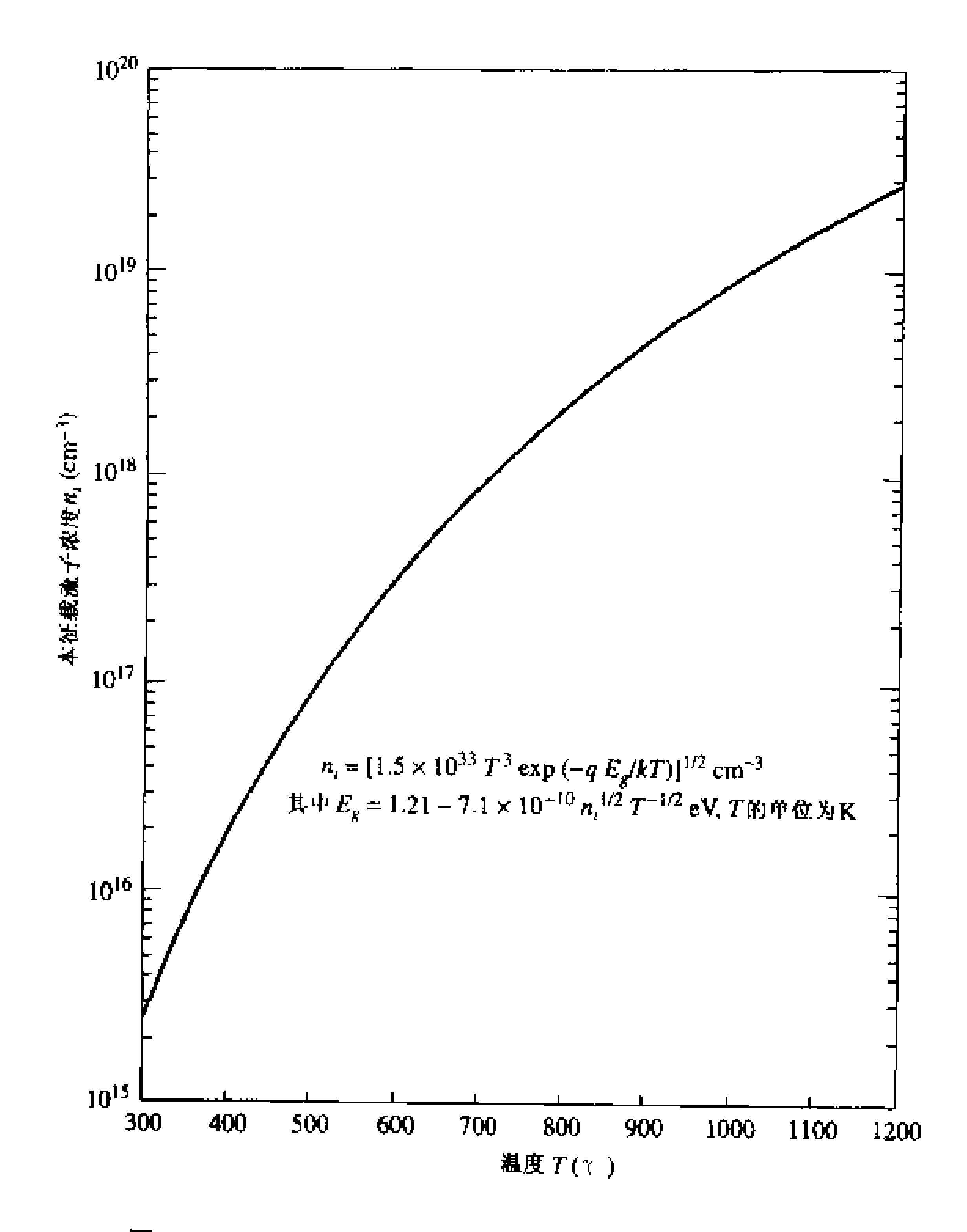


图 2.10 300℃～1000℃时，硅中的本征载流子浓度 n_i[10]

抛物线速率系数 B 取决于 SiO_2 的性质，因此与硅中的点缺陷没有直接关系。然而，B 确实与氧化剂在二氧化硅中的扩散系数有关。因为某些硅中的杂质在氧化过程中会进入二氧化硅层并破坏它的结构，扩散系数（因而抛物线速率系数 B）也会增加。但是，这个增加与线性速率系数 B/A 的增加相比很小。因此，对较厚的二氧化硅，硅中高的杂质浓度引起的氧化速率增强量比较小。

掺氯氧化[†] 当氧化剂中含有氯时氧化速率也会增加（一般增加 10%～20%）。氯通常来自 HCl，Cl_2 或有机化合物，进入到 Si-SiO_2 界面附近可提高器件的电学特性，尤其是第 8 章讨论的 MOS 器件。通过在氧化剂中加氯的方法改善 Si-SiO_2 系统的性质必须非常小心。当浓度太

低时,性能的提高非常有限,而浓度过高时会腐蚀硅表面,或在 Si-SiO_2 界面形成气泡使氧化层开裂。

氮氧化硅 尽管适当厚度的氧化硅可阻止杂质原子的扩散,但是杂质原子,尤其是硼可能穿透很薄的氧化层(<5nm 厚)。

当 MOS 晶体管的栅氧变得更薄时,这种有害的杂质迁移会改变晶体管的电学性质。如果对生成的二氧化硅在含氮的气氛如氨气(NH_3)中热处理,会在氧化物中增加部分氮的含量,能阻止杂质扩散。在氧化层的顶部(也可能是底部)形成的富氮层,增加了对杂质穿透的阻碍作用。

高介电常数氧化层 MOS 晶体管或存储单元的栅电极与沟道或存储区间需要更好的电容耦合,氧化硅或氮化硅已不能提供所需的电容,需要使用更高介电常数的绝缘体。氧化钽(Ta_2O_5)等材料能满足要求,氧化锆和氧化铪也是有希望的候选者。也有人提出采用更高介电常数材料如钛酸钡锶(BST)。为了能利用这些介电常数更高的氧化物,就必须解决与电极紧密接触的问题。由于二氧化硅非常稳定,通常在硅电极和高介电常数氧化物间加一薄层的 SiO_2。相关的寄生串联电容使得总电容比只有高介电常数氧化物时低一些。

2.4 光刻和图形转移

光刻 在硅片上形成 SiO_2 保护层后,必须选择性地去除部分区域,以便进行掺杂。选择性刻蚀离不开称做光刻胶的光敏聚合物材料。首先,在高速旋转的氧化硅片上滴一些液态的光刻胶。等胶烘干后,将由透明和不透明区域组成的玻璃板(称做掩膜版或光刻版)放在硅片上,如图 2.11(a)所示,用显微镜对准。然后,用紫外光对光刻胶曝光以改变其结构。对正胶,曝光处的分子键被打破,而负胶的分子键在曝光后是交联(聚合)的。光刻胶中弱的键合区域或未聚合的区域被选择性地在溶剂中溶解掉,因此未被溶解的、耐酸性的、变硬的胶层将掩膜版上的图形复制在 SiO_2 上(图 2.11(b))。类似地,也可在 IC 工艺中的其他材料层的表面形成光刻胶图形。

在大部分直接加工的硅平面工艺中,硅片上所有电路单元的图形同时曝光。然而,当器件尺寸变小时,不仅必须分辨的最小特征尺寸减小了,各层图形之间的对准也要变得更加精确。在热循环过程中,热应力、掺杂原子引起的应力或其他材料层的引入会造成芯片轻微的形变,这样,后面的掩膜版将不能非常精确地与前面形成的图形对准。这种形变会限制不同掩蔽层之间的对准精度。一种解决的办法是采用步进式光刻机(图 2.11(c)),它一次只对晶圆片上的一个单元曝光。完成一个单元的曝光后,晶圆片移动或步进到下一个单元继续曝光。尽管这一工艺的机械结构更复杂,而且比整个晶圆片曝光速度慢,但是却能够提高对准精度。另外,对包含单个重复单元的大掩膜版上的图形进行缩小(大约 5 至 10 倍),可以获得更小的器件特征尺寸。

先进的光刻技术 集成电路中要包含有更多的晶体管,由光刻工艺决定的最小特征尺寸必须不断地减小。在直接曝光工艺中,如果光的波长接近于特征尺寸,就会发生衍射效应,因而光的波长会直接限制特征尺寸的减小。由于衍射效应,硅片表面的电场和光强会在一段距离内逐渐变化,这个变化范围与曝光波长有关(图 2.12(a)),因而难以制造出陡峭的边界。因此,采用短波长曝光可得到更小尺寸的图形。为了保证曝光强度,通常采用汞灯。汞灯在 UV

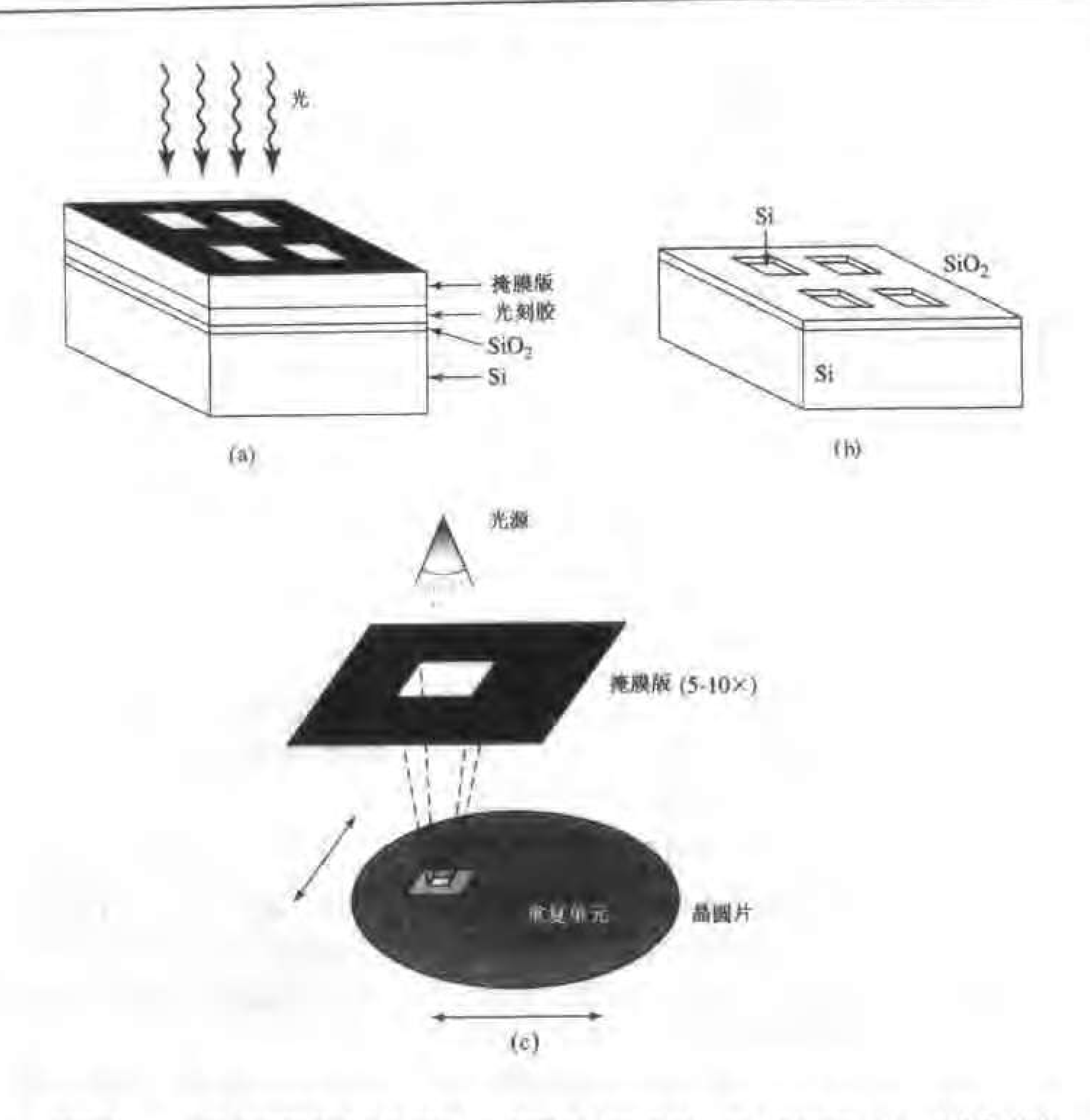


图 2.11 (a)通过掩膜版对光刻胶曝光,确定出被刻蚀的二氧化硅区域;(b)坚膜后的光刻胶保护层下面的二氧化硅不被化学去除;(c)在步进式投影光刻系统中,光线通过掩膜版上的图形。每一图形成的像被缩小聚焦在硅片上的一个单元,一个单元上的所有图形是同时曝光的。然后,硅片被移动(步进)到下一个单元,重复进行曝光

光区有三条强的谱线,波长分别为 436nm(G 线)、405nm(H 线)和 365nm(I 线)。最精细的图形可采用短波长的 I 线而不是 G 线曝光。更短波长的光源来自于激光,如 KrF 或 ArF 激光器;这些激光光源常用的波长为 248nm 和 193nm,使用 F_2源的 157nm 的光源正在开发。要想进一步减小波长很困难,因为用于制作透镜和掩膜版的大部分材料在短波长下变得不透明。因此必须采用反射式光学系统,这使光刻机的设计和使用复杂化。然而,人们还是花费了很大精力去发展超短波长如极紫外(EUV)曝光系统,其工作波长在 13nm(位于软 X 射线区域),这时仍然可以在相当大的曝光区域内(通常是一个或多个芯片)同时对各种图形曝光。除了研究复杂的光学系统外,获得短波长下的高强度光源也是一个挑战。

除了采用更短的波长外,用芯片上更复杂的光刻胶加工工艺也可制备出更小的图形。光刻胶下的抗反射膜可减小多次反射,而采用移相掩膜可获得更陡峭的图形边界。如图 2.12(a)所示,在精细图形的边缘附近光强逐渐发生变化,如果两个图形间距很近时中间区域的光强不再为零(图 2.12(b))。然而,光强与光波的电场强度的平方成正比。如果两个图形上的

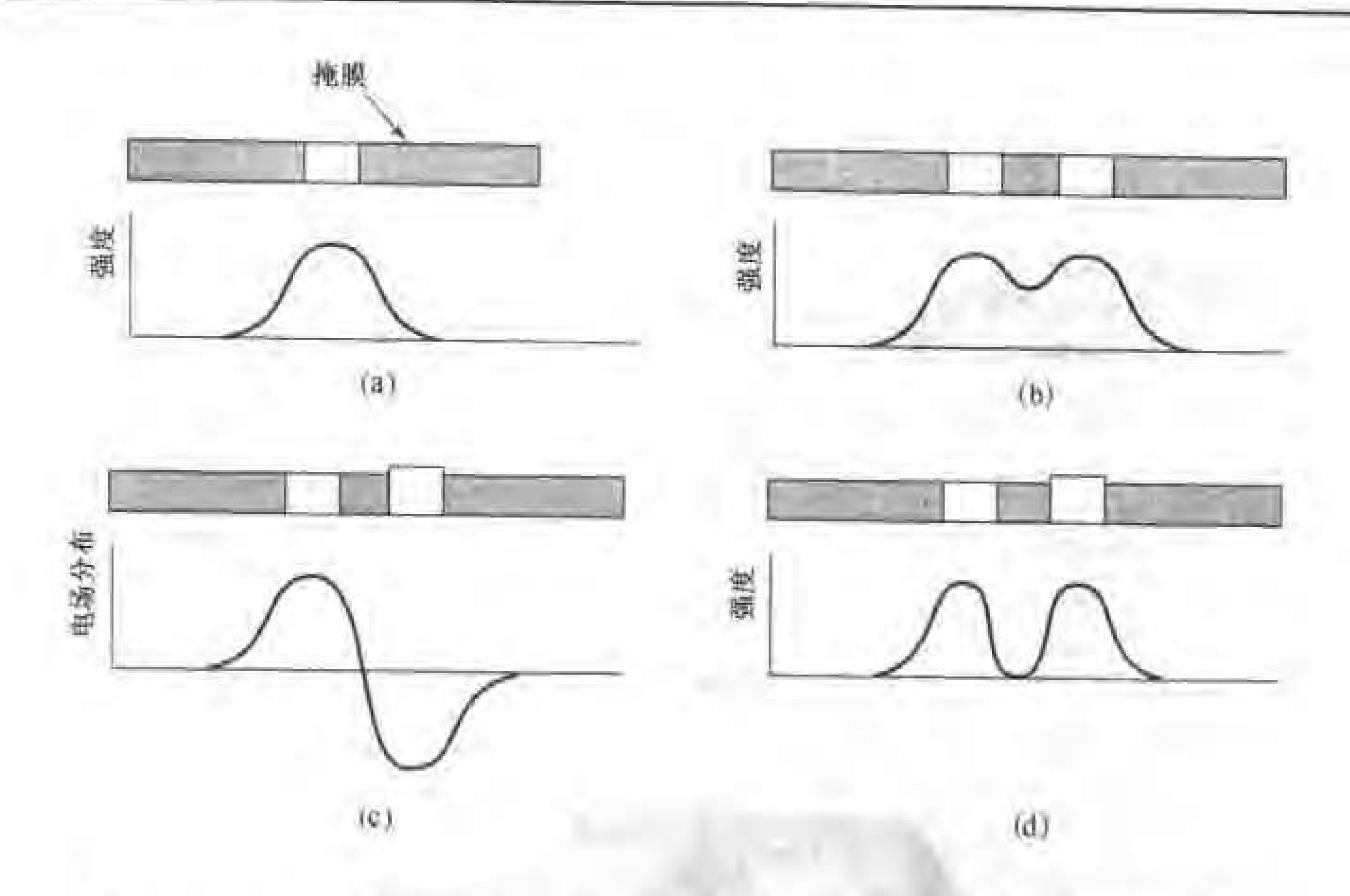


图 2.12　(a)由于衍射效应，在精细图形的边缘附近光强逐渐变化；(b)两个图形间距很近时中间的光强不再为零；(c)改变通过掩膜的路径长度可使电场位相移动 180°；(d)最终使得图形间的光强变为零

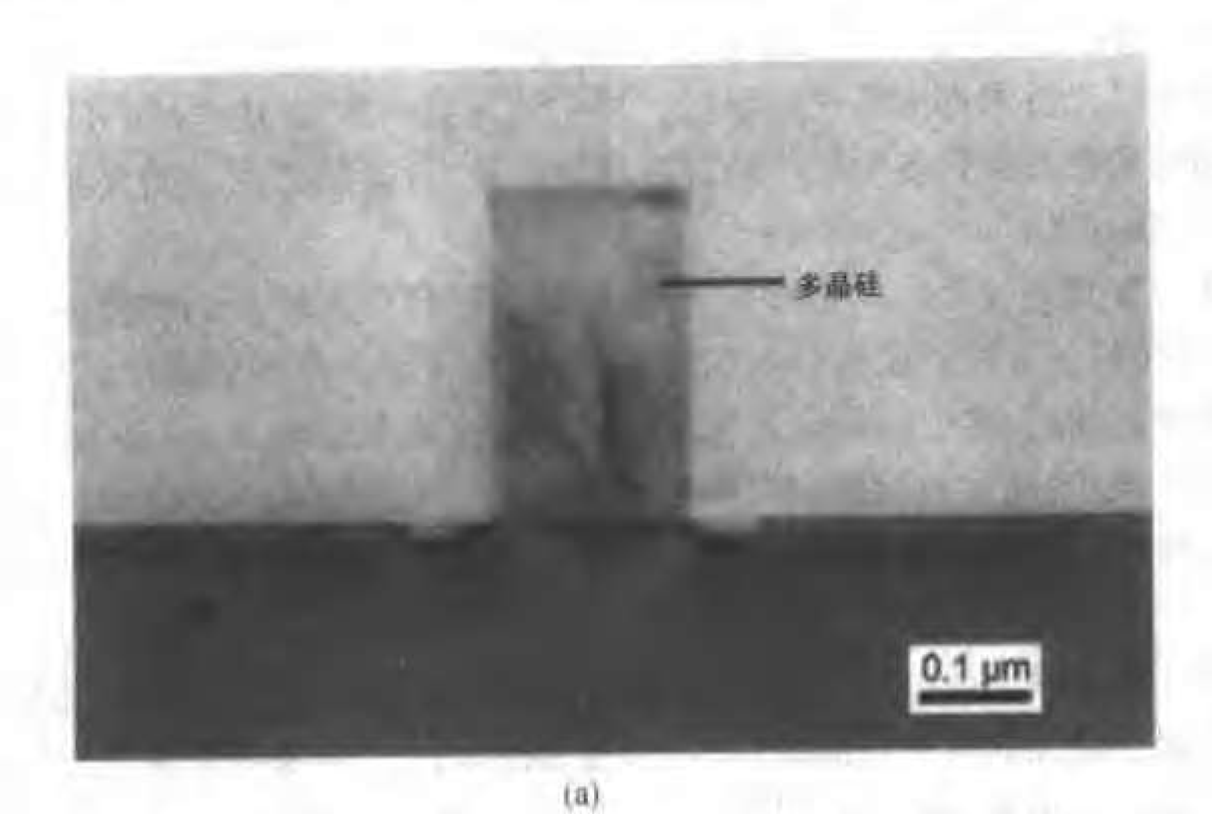
电场符号不同，则两个图形中间某些点的电场一定为零(图 2.12(c))，它的平方(光强)也必定为零(图 2.12(d))。掩膜版中部分区域加厚，使光线通过某些透光区的光程增加半个波长，相应的电场相位改变 180°，从而电场符号发生改变，而通过相邻图形的电场符号不变。这样通过两个相邻图形的电场符号相反。也可应用非线性对比度增强技术，稍微改变图形边缘附近的光强就可使光刻胶的化学性质发生较大的变化。

细线条的均匀性会受到 IC 芯片表面不平坦的限制。由于曝光光束聚焦景深的限制，与前道工序形成的高台阶相交叉的线条可能不好限定。如果采用两层或三层的多层光刻胶结构可以解决这个问题。第一层厚的光刻胶填满芯片上的沟槽部分，形成更均匀的表平面。在上面的薄光刻胶中曝光形成精细图形，然后通过定向刻蚀将图形转移到下面厚的光刻胶上。

将更短的光源波长和更复杂的光刻胶系统结合起来，光学光刻得到的特征尺寸可减小到 130nm。光学光刻的主要优点是保持了在相当大的面积上能同时对所有图形曝光。图 2.13(a)是光学光刻和各向异性刻蚀形成的横向尺寸约为 180nm 的多晶硅栅的截面图。

采用不同于光学光刻的其他方法，如利用电子束、X 射线或离子束对光刻胶进行曝光可得到更小的图形。在电子束光刻系统中，聚焦的电子束将能量转移给光刻胶使其曝光。与光学光刻不同的是，它不能对所有复杂的图形同时曝光，而是要按顺序偏转电子束来曝光各图形单元。对电子束的控制信息储存在计算机中，这种光刻方法不需要掩膜版。电子束光刻最常用于制备几何尺寸很小的掩膜版，然后将掩膜版用在常规的光刻曝光技术中。电子束可被聚焦成远小于最小特征尺寸小的束斑，然后使其移动通过各单元，每一图形由电子束多次扫描完成。电子束可被聚焦成矩形束斑，像堆积木一样重复曝光形成图形。无论哪种情况，电子束顺序曝光的特性限制了芯片的加工速率。但是，曝光一个硅片所增加的时间，至少部分地被得到更高密度的电路所补偿。

X 射线光刻使用 X 射线通过掩膜版对光刻胶进行曝光。X 射线光刻与光学光刻一样，能



(a)

(b)

图2.13 (a)长度约为180nm的多晶硅栅，栅氧及硅衬底中浅结的横截面透射电镜图(引自Accurel Systems International Corp.)；(b)宽度为500nm间隔为1.5μm由各项异性刻蚀形成的线条。光刻胶覆盖了180nm $TaSi_2$ 和260nm多晶硅双层结构。注意垂直于表面各层的均匀性(引自D. Dorda, Siemens Corporation)

同时对许多图形曝光，但是短波长X射线可以形成更精细的图形。X射线光刻不如光学光刻和电子束光刻发展得成熟，主要是受到X射线光源和掩膜版的限制。大部分常规的X射线源是点光源，当穿过掩膜版和硅片时会发散，因而限制了掩膜版与硅片之间的距离。同步加速器可以产生部分校准的X射线束，是可能的但是非常昂贵的高强度X射线源。

X射线掩膜版也很难制备。因为X射线必须容易穿透掩膜版上薄的透明区域，这非常困

难。使用 X 射线遇到的更基本的问题是一次曝光的面积会受到由加工引起的芯片形变的限制。任何大面积曝光技术都面临这一限制。与光学光刻一样，即使硅片上只有小的形变存在，对一个单元的曝光会影响到各层图形的对准精度。最后，必须考虑高能 X 射线对硅片上器件有源区可能造成的损伤。

图形转移 在光刻胶上形成图形后，SiO_2或其他材料中未被光刻胶保护的区域要被刻蚀掉，才能将图形转移到芯片。要得到 SiO_2的图形，可在含有氢氟酸的腐蚀液中腐蚀掉 SiO_2，使硅表面裸露出来。然后，再去掉剩余氧化硅区域上保护其不被腐蚀的光刻胶。这时，硅片上的部分区域被 SiO_2保护，而在氧化物窗口中的裸硅(图 2.11)用于随后的掺杂。有很多种的液态化学溶剂可选择性地腐蚀某些材料而不损坏下面的物质。这种高选择性是液态或湿法腐蚀的一个优势。

干法刻蚀 随着 IC 单元上需要定义的尺寸的减小，湿法腐蚀工艺明显受到几个限制。一个主要问题是湿法腐蚀通常是各向同性的，垂直于硅表面的腐蚀与掩蔽膜下的横向腐蚀同时进行(图 2.14(a))。因此，腐蚀出的图形总是比掩蔽膜上的图形大。干法(等离子或反应离子)刻蚀技术可以是各向异性的，从而解决了这个难题(图 2.14(b))。(反应离子刻蚀是等离子刻蚀的一种特殊情况，可以通过优化工艺条件获得高度各向异性的图形。)

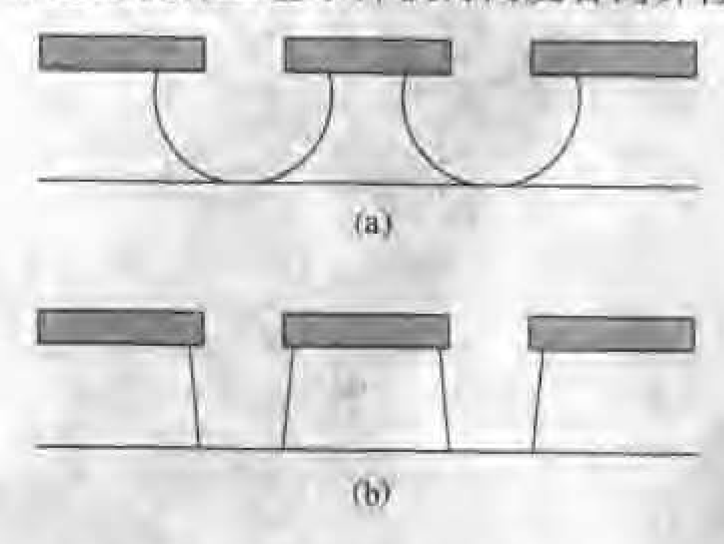


图 2.14 (a)在各向同性的湿法腐蚀或干法刻蚀中，化学反应起主导作用，引起掩蔽膜下的严重钻蚀；(b)各向异性的离子辅助干法刻蚀产生近乎垂直的图形，能保持掩蔽膜的尺寸

等离子体是在高频电场中产生的，由处于激发态的中性粒子(原子团)、离子和电子组成的近乎电中性的混合体。产生激发态的中性粒子需要的能量比将分子电离需要的能量低，因此原子团的数目远大于离子的数目。尽管等离子体本身是电中性的，但等离子区通常与晶圆片表面有一定的距离，因此它们之间的电场能加速离子向表面运动。

干法刻蚀选择能与被刻蚀材料发生化学反应的气体作为刻蚀气体。选择刻蚀气体的一个主要考虑是反应产物必须是挥发性的。硅和硅的化合物可被含氟的气体有效刻蚀，而铝用含氯的气体去除。铜的刻蚀产物在室温下通常是不挥发的，这使得铜很难刻蚀；通常选择所谓的"大马士革镶嵌"(Damascene)工艺来形成铜的图形，这将在 2.7 节中介绍。有机光刻胶一般在氧等离子体中干法刻蚀，反应产物是水蒸气和二氧化碳，通常还要加入不参与反应的稀释气体来帮助气体的流动。

在干法刻蚀中，带掩膜的硅片放置在等离子体中。处于激发态的中性原子团与被刻蚀材料在其暴露的区域发生化学反应，等离子体中的离子，轰击表面以去除暴露在表面的材料。垂

直方向的刻蚀速率大于单独化学去除和物理去除的速率之和。离子轰击水平表面有助于反应产物从表面解析，使新鲜的反应剂更容易到达表面而增加反应速率。另外，激发态的原子团轰击水平表面可产生晶格损伤，削弱化学键，使化学去除进行得更容易。当表面处离子与受激中性原子团的比例增加时，即所谓的"高密度等离子体"，则材料表面的物理溅射变得很重要。

通过适当选择反应气体，电场和反应器几何形状，刻蚀反应有可能是各向异性的，这样在掩膜版边缘被刻蚀材料中可形成几乎垂直的侧壁，如图2.14(b)。在定向干法刻蚀时，反应物、部分分解物及反应产物可能会淀积在刻蚀图形的侧壁。只有少量的离子能够轰击到几乎垂直的表面，所以这些淀积物保留在表面，阻止了横向刻蚀，从而得到高度的各向异性的图形。特定的电场方向和低的气压(因而离子散射弱)能够增强垂直方向的离子轰击，因而加强了各向异性刻蚀。通过改变化学去除和离子辅助去除的重要性，就可改变各向异性刻蚀的程度。

尽管高度各向异性刻蚀是可能的，但是干法刻蚀工艺的化学选择性却不如湿法腐蚀大，而且选择性会随各向异性刻蚀程度的提高(即离子轰击变得更为重要时)而下降。在刻蚀过程中掩膜材料会受到严重的侵袭。有限的选择性会造成对被刻蚀材料下面的衬底的刻蚀；因此，干法刻蚀必须在刻蚀完成后立刻停止。这一过程可以通过监控反应中的特征光谱来进行。另外，随着被刻蚀材料厚度的不断减小，由于光学干涉效应，可以观测到照射在硅片上的激光束的反射光的变化。当材料被完全刻蚀后，反射光的特性发生改变，需要的过刻蚀完成后就可以停止刻蚀。

对多层材料的刻蚀具有很大的挑战性。由于不同材料具有不同的各向异性刻蚀特件，可能会造成凹槽和钻蚀，使后面的工艺复杂。图2.13(b)给出了刻穿不同材料层后均匀的形貌。

在进行干法刻蚀时，还必须考虑高能粒子环境的影响；反应室中激发态的离子和高能光子轰击硅表面，可能损坏正在制作的器件。等离子增强工艺将在2.6节中进一步讨论，那里将介绍等离子化学气相淀积。

2.5 掺杂和扩散

掺入硅中的杂质原子的分布一般由两步工序决定。首先，通过离子注入、气相淀积或在硅表面涂敷含有掺杂剂的涂层而将杂质原子引入硅表面。其次，进行推进扩散使杂质原子在硅片中重新分布。最后的杂质分布主要由硅表面杂质的初始状态决定，而扩散深度主要取决于推进扩散的温度和时间。因为半导体器件的特性强烈地依赖于杂质分布，下面将对掺杂工艺进行详细的讨论。

2.5.1 离子注入

离子注入是在半导体中引入杂质原子的一种可控性很强的方法。被注入的杂质原子首先被离化，然后通过电场加速获得高能量(典型值为25～200keV)。这些高能离子束轰击半导体表面(图2.15)，进入暴露的硅表面区域。掩蔽材料可以是氧化硅或IC本身结构中的其他材料。因为离子注入对硅片的加热不严重，光刻胶仍可用做芯片上选择性离子注入的掩蔽膜。更高密度的材料对注入的阻挡作用更有效，采用的掩蔽膜可以更薄。

离子的穿透深度通常小于1μm，离子注入时晶体会受到严重的损伤。因此必须用退火工艺来修复晶格的损伤，以保证注入的杂质原子替代硅的位置，作为施主或受主。离子注入后，

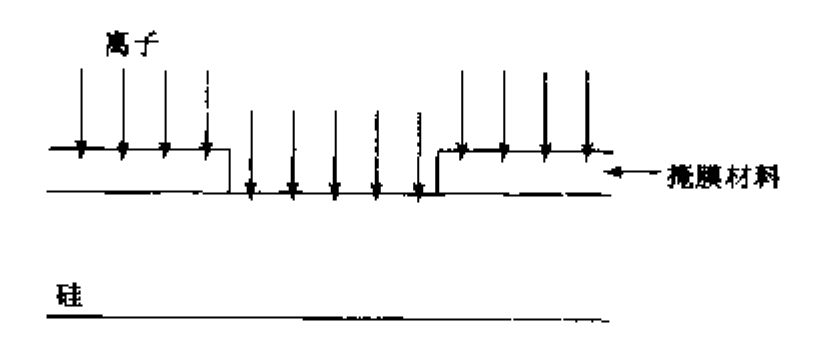


图2.15 在离子注入中,高能离子束轰击半导体表面上的选择性区域,进入到暴露的硅片中

可根据需要通过随后的扩散进行离子的重新分布。

离子注入能够精确控制进入硅片中单位面积的杂质原子数。因为杂质是作为带电离子被注入的,在注入过程中可以在离子通道的位置放置一个比较简单的电荷检测装置对离子进行计数。当达到所需数目时,离子束就被关断。

通过精确的控制,进入硅片的掺杂剂量(杂质原子数/cm^2)分布在大约$5\times10^{11}cm^{-2}$到大于$10^{16}cm^{-2}$。最低剂量的离子注入用于精确调节器件特性,而非常高剂量的注入用于形成低阻欧姆接触。离子注入浓度的上限与气相淀积可得到的浓度相当。注入机中典型的离子束流为1mA量级,对应于单价离子流为$6.25\times10^{15}s^{-1}$。对高剂量注入,注入机可产生非常高的电流以减小注入时间。

除了能精确控制总的剂量,离子注入的掺杂物纯度也非常高。高纯度的注入是这样得到的:在离子源附近放置质谱仪对离子进行挑选,只有所需要的离子才能到达硅片。

由于注入的离子能够穿透晶圆片表面,因此也可利用离子注入工艺透过晶圆片表面的另一层材料(例如SiO_2)向硅中引入杂质原子。这样,经过高温热循环形成热氧化物后再进行离子注入,杂质原子的分布只由注入能量决定,而不是由长的热氧化时间内的扩散决定的。只需要适当的热过程就可去除注入损伤和激活杂质。因此,杂质原子的横向扩散也减小了。然而,当离子束穿过氧化层时,会将能量和动量转移给氧原子。某些氧原子会被推进(碰撞)到下面的硅中,从而会影响器件的工作。

在非晶层中,注入离子的分布接近于Gauss分布(即分布满足$A\exp-(x/\lambda)^2$)。Gauss分布在表面下的平均穿透深度处有一极大值称为投影射程R_p。Gauss分布的宽度用投影射程的标准偏差ΔR_p表示。总的分布与表面下的深度x之间的关系为

$$C(x)=C_p\exp\left[-\frac{(x-R_p)^2}{2\,\Delta R_p^2}\right] \tag{2.5.1}$$

其中峰值杂质浓度(原子数/cm^3)C_p与注入剂量N'(原子数/cm^2)之间的关系为

$$C_p=\frac{N'}{\sqrt{\pi}(\sqrt{2}\,\Delta R_p)} \tag{2.5.2}$$

从等式(2.5.1)可看出,$\sqrt{2}\Delta R_p$是描述注入离子在空间展宽的特征长度。文献中报道了许多关于R_p和ΔR_p的实验结果[11]。三种最常用的杂质原子在硅中对应的参数示于图2.16至2.18中。关于一种离子在多层结构中的穿透问题,可通过考虑每一层中的能量损失来近似处理。

等式(2.5.1)和(2.5.2)只适用于离子注入进入非晶材料中的情况,这时离子散射可看做

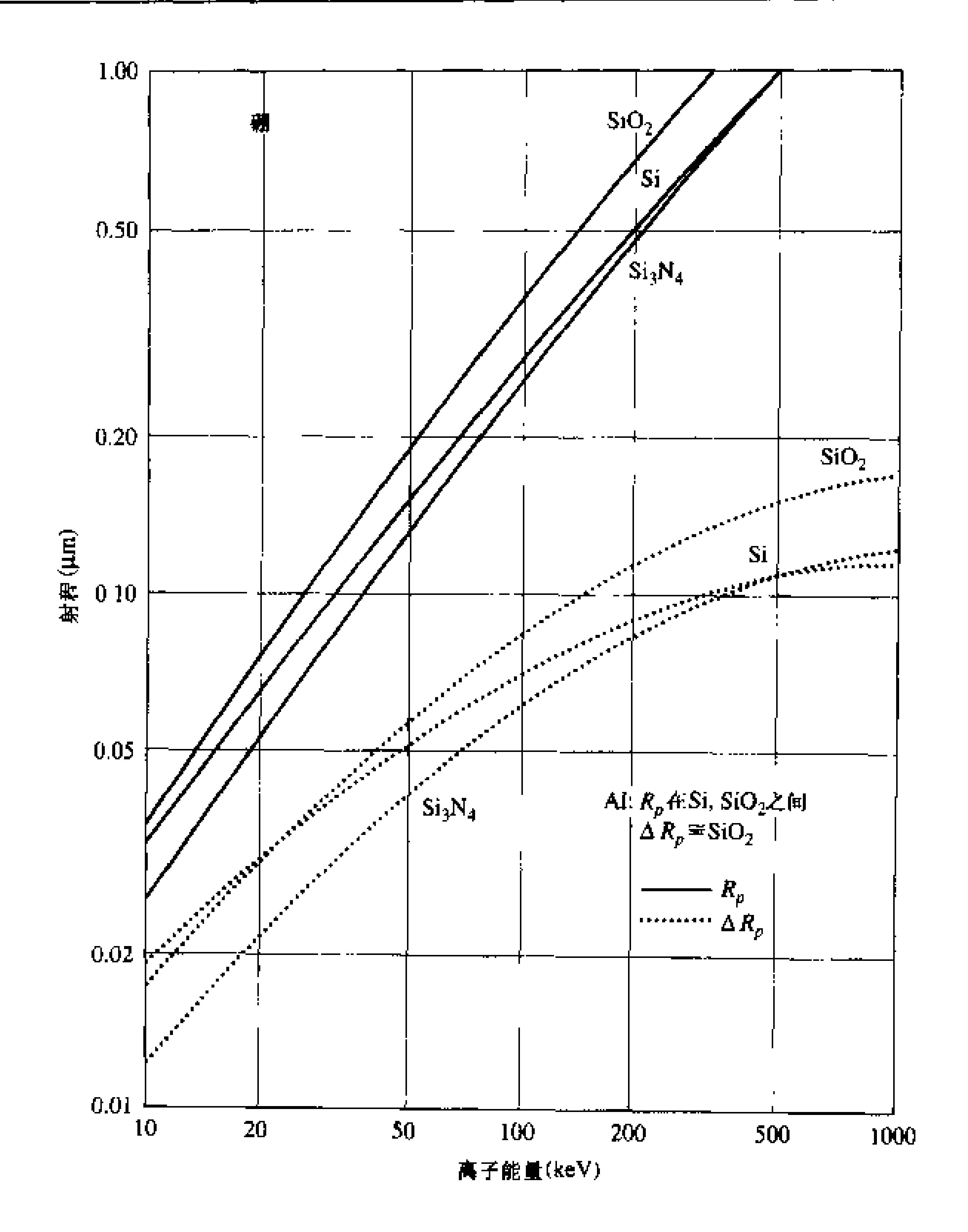


图2.16　硼注入Si、SiO_2、Si_3N_4和Al中的投影射程R_p和标准偏差ΔR_p[11]

是各向同性的。在晶体材料中，如果离子是沿着原子间距较大的方向（沟道方向）进入晶格的，离子的穿透深度可能远大于R_p。对大部分IC工艺都不希望出现沟道效应，由于沟道效应对离子束在硅片表面的入射角敏感，如果将硅片倾斜，使得离子束和晶轴间夹角为7°或者在硅片表面覆盖一层非晶SiO_2就可减小沟道效应。无论采取两种措施中的哪一种，都可以使用非晶材料的射程参数。多晶材料中有多种晶向，因而不能通过倾斜硅片来避免沟道效应。

离子注入的掩蔽膜可采用厚的SiO_2层，也可以采用金属或有机光刻胶，因为在注入过程中硅片温度的升高是有限的。使用光刻胶做掩蔽膜很方便，但是必须注意，要确保注入过程对有机材料的加热没有使其流动或变形。对大束流注入，这是十分有害的，这时注入设备必须要将注入时硅片上的热量散去。注入离子束的能量也可改变光刻胶的结构，使其不能在后面的湿法腐蚀中很容易地去除。在这种情况下，必须在氧等离子体中去除光刻胶。

由于能够精确控制注入的杂质原子的剂量、纯度和位置，离子注入已成为最主要的集成电路选择性掺杂的方法。第9章中的进一步讨论表明，离子注入是先进的MOS工艺中的一个基本组成部分。

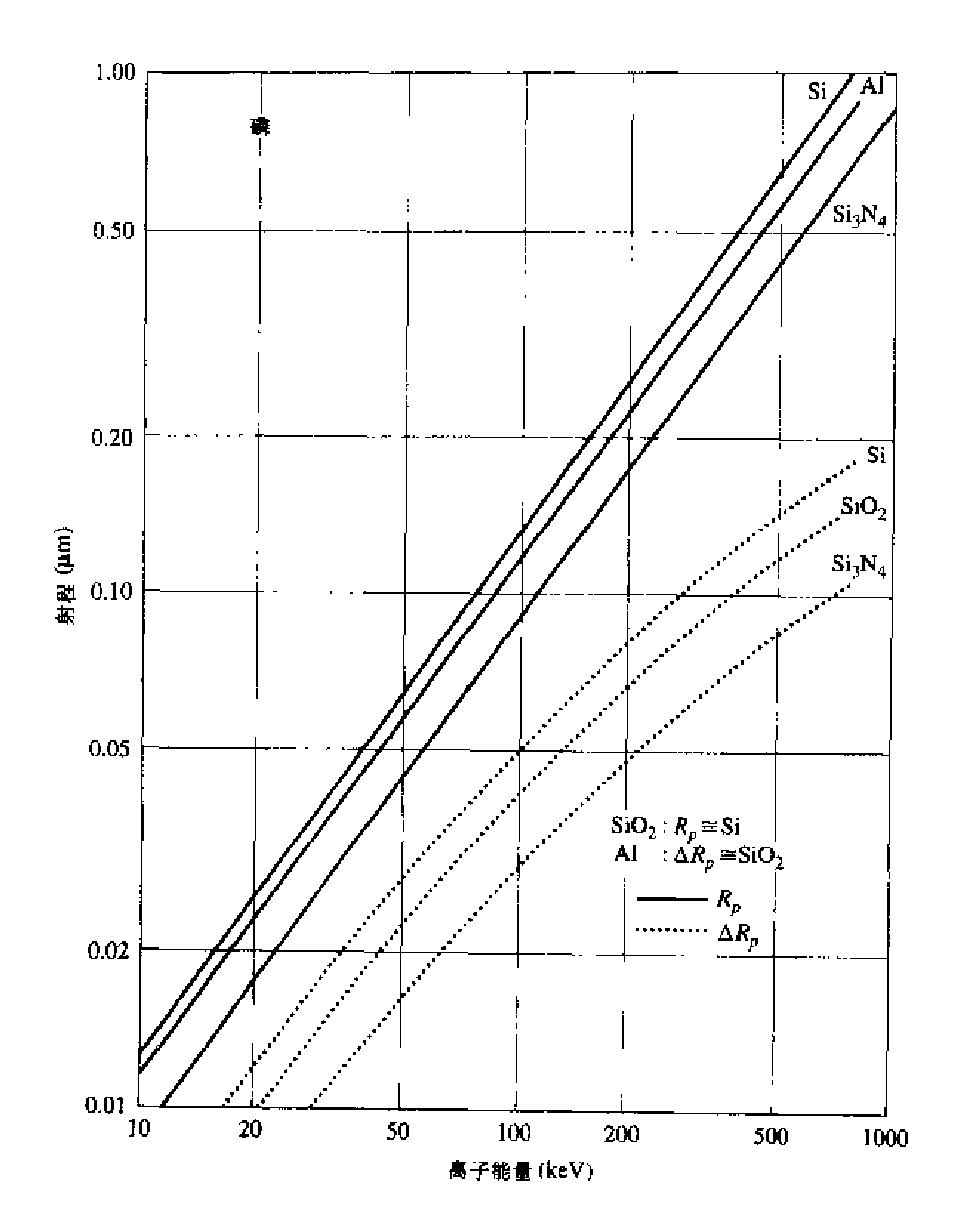


图 2.17 磷注入 Si、SiO_2、Si_3N_4 和 Al 中的投影射程 R_p 和标准偏差 ΔR_p[1]

2.5.2 扩散

硅中的杂质原子只要具有足够的能量就可以在晶体中迁移。它们可以从初始淀积的高浓度区向晶圆片深处的低浓度区扩散。杂质原子的扩散类似于 1.2 节讨论的自由载流子的扩散。两者主要的差别在于造成明显移动所需要的温度不同。一般来讲，晶格中必须有点缺陷(通常是硅空位或间隙硅原子)存在时，杂质原子才能在晶格中移动，而价带或导带中的自由载流子不需要与点缺陷相互作用就可以移动。另外，只有当温度达到 800℃~1000℃时，才会发生显著的杂质原子的扩散运动。

在距表面深度为 x 的一小区域 dx 内(图 2.19)，杂质原子的浓度随时间的变化可写成单位面积从左边流入和从右边流出这一小区域的杂质原子流量之差

$$\frac{\partial C}{\partial t}dx = F(x) - F(x + dx) \tag{2.5.3}$$

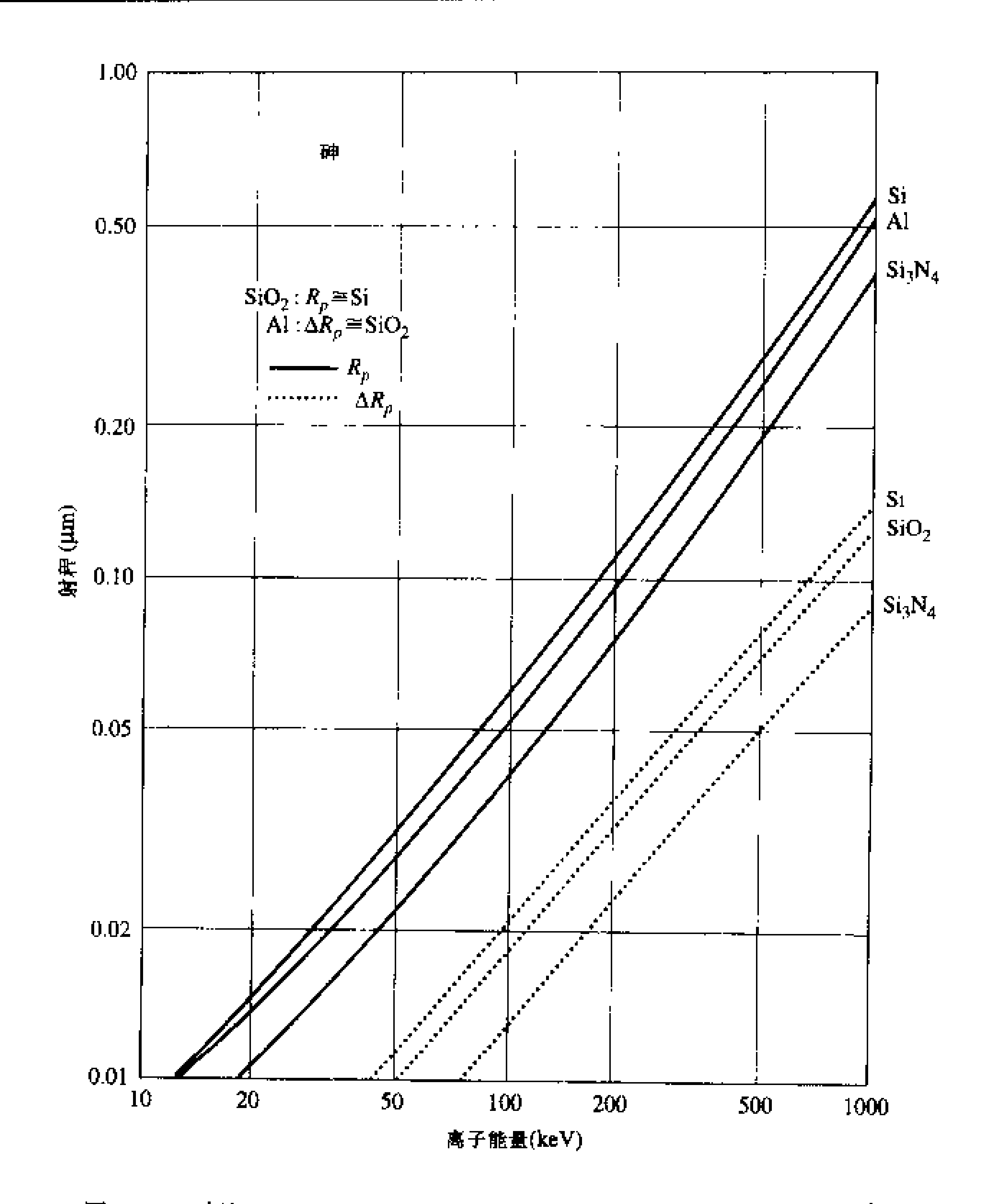


图 2.18 砷注入 Si、SiO_2、Si_3N_4 和 Al 中的投影射程 R_p 和标准偏差 ΔR_p[1]

式中最后一项可用泰勒级数展开，取前两项

$$F(x + \mathrm{d}x) \approx F(x) + \left(\frac{\partial F}{\partial x}\right)\mathrm{d}x \tag{2.5.4}$$

可以得到

$$\frac{\partial C(x)}{\partial t} = -\frac{\partial F}{\partial x} \tag{2.5.5}$$

与 1.2 节中看到的一样，在一级近似下，扩散流正比于浓度梯度

$$F = -D\frac{\partial C}{\partial x} \tag{2.5.6}$$

参数 D 称为扩散系数，描述的是杂质原子在晶格中移动的难易程度，它强烈地依赖于温度。几种常用杂质的扩散系数如图 2.20 所示。这里考虑的简单模型忽略了扩散的许多效应，这些效应在更高的掺杂浓度下才变得重要。因此，图 2.20 中给出的值只在低、中等掺杂浓度下才

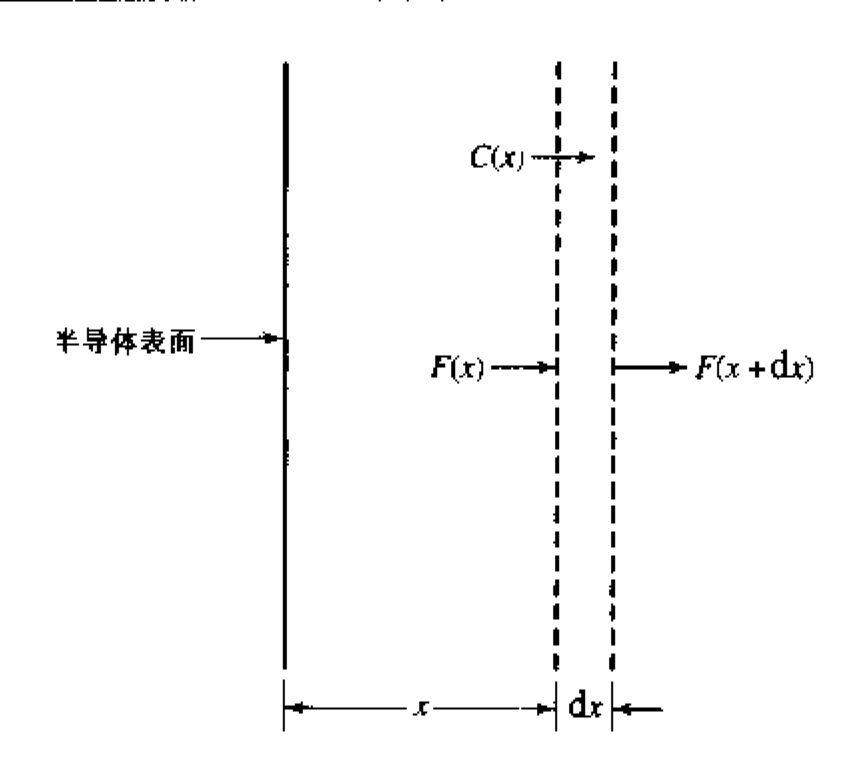


图 2.19 在 dx 区域内杂质浓度的增加与进入这一区域的原子净流量有关：$F(x)-F(x+dx)$

适用。当杂质浓度增高($\geq n_i$)时必须小心使用，因为那时扩散系数本身也是杂质浓度的函数。在这一问题上，杂质 P 尤其棘手。关于二级效应的详细情况在后面讨论。结合等式(2.5.5)和(2.5.6)有

$$\frac{\partial C}{\partial t}=D\frac{\partial^2 C}{\partial x^2} \tag{2.5.7}$$

等式(2.5.7)(有时称为 Fick 第二定律)可解出 $C(x,t)$ 的解析表达式。

考虑相关的边界条件，等式(2.5.7)可得到具体情况下的解。根据半导体工艺中最常用的扩散条件，等式(2.5.7)有两种不同的边界条件，可得到两种形式的解。一种是固定表面杂质浓度(原子数/cm^3)的扩散，得到的是余误差分布函数，而固定扩散原子总数(原子数/cm^2)的扩散，得到的是 Gauss 分布函数。这两种分布如图 2.21 所示。

气相淀积 用气相淀积源在半导体中引入杂质原子时，图形化后的硅片放置在类似于氧化时用到的扩散炉中，含有所需杂质(通常为磷或硼)的气体从炉中通过。杂质进入硅片的数量只能小于或等于炉温下杂质在硅中的固溶度。

控制进入硅中杂质原子数量的最好方法是调节气流，使得硅表面的杂质浓度达到它的固溶度。几种常用掺杂元素在硅中的固溶度如图 2.22 所示。

为了将引入的杂质原子数限制在能使器件正确工作所要求的值，气相淀积通常采用相对低的温度和比较短的时间。在这一淀积过程中，杂质原子的穿透深度一般比较小，需要随后的推进扩散使得淀积的原子重新分布达到需要的深度。

气相淀积时，硅表面暴露在固定的杂质浓度下，等式(2.5.7)相应的解表明，淀积后杂质原子沿 x 方向(测量离开表面的尺寸)为余误差函数分布。

$$C(x,t)=C_s\operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right)=\frac{2C_s}{\sqrt{\pi}}\int_{x/2\sqrt{Dt}}^{\infty}\exp(-v^2)\,dv \tag{2.5.8}$$

其中，C_s 为杂质原子的表面浓度。注意到

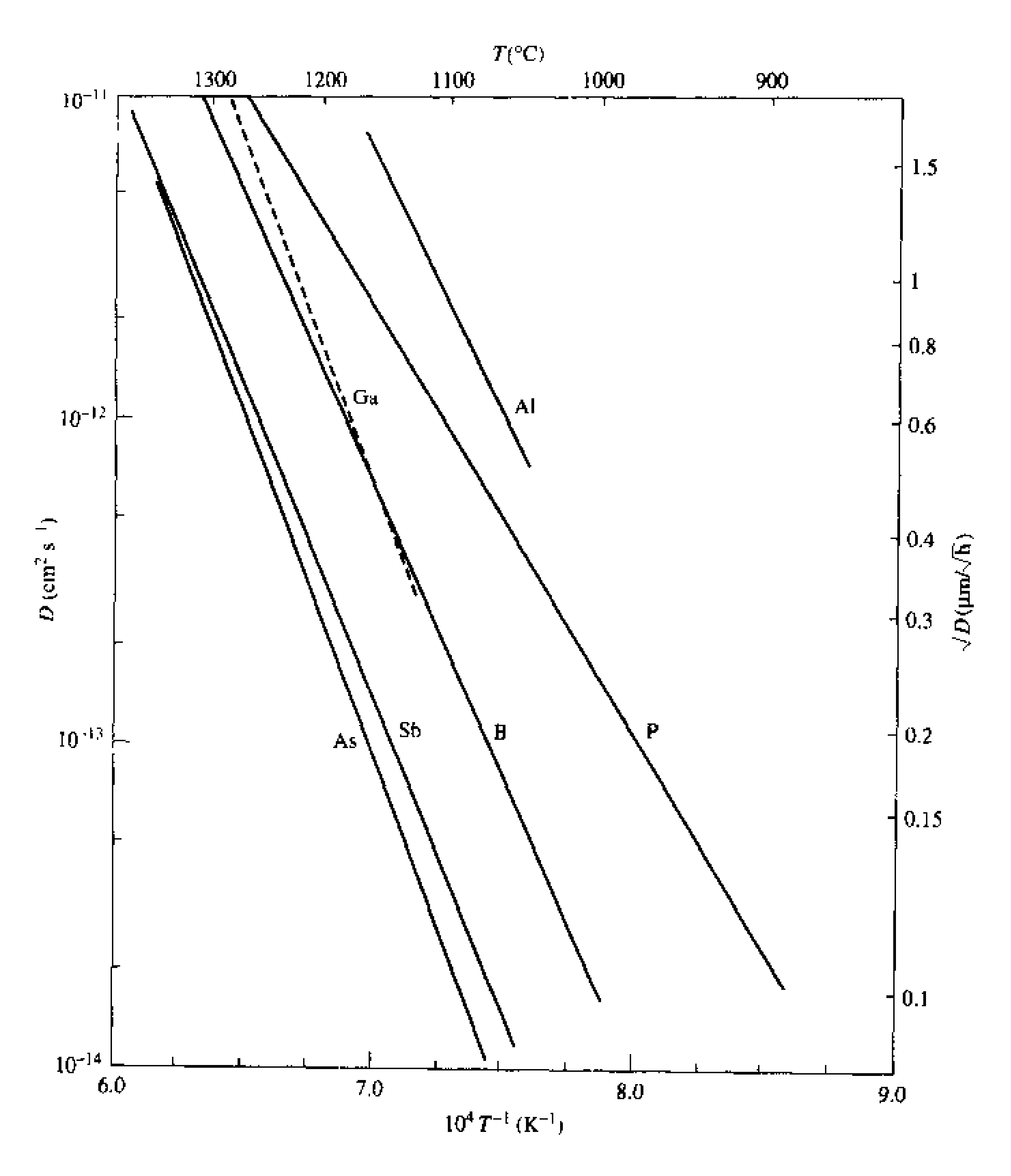


图2.20 常用杂质在硅中的扩散系数(低浓度下)与温度的关系[12]

$$\mathrm{erfc}(\eta) = 1 - \mathrm{erf}(\eta) = 1 - \frac{2}{\sqrt{\pi}}\int_0^{\eta}\exp(-v^2)\,\mathrm{d}v \tag{2.5.9}$$

因此,erfc(0)=1,余误差函数随着自变量 η 的增加而迅速下降(图2.21)。

只要硅表面的杂质浓度为常数[$C(0,t)=C_s$],不管以何种方式引入的杂质,等式(2.5.7)的解都是余误差函数。图2.21中用于归一化 x 轴的参数 $2\sqrt{Dt}$,表示与特定扩散过程有关的特征扩散长度 L,它描述了杂质的穿透深度。注意到扩散深度只随扩散时间 t 的平方根而增加。

如果已知表面浓度、固溶度、扩散系数和扩散时间,可以从等式(2.5.8)中计算出杂质的分布。更加重要的是,还可以计算出由扩散引入的单位表面总的杂质原子密度 N',这个值等于余误差分布函数 $C(x,t)$ 对 x 的积分

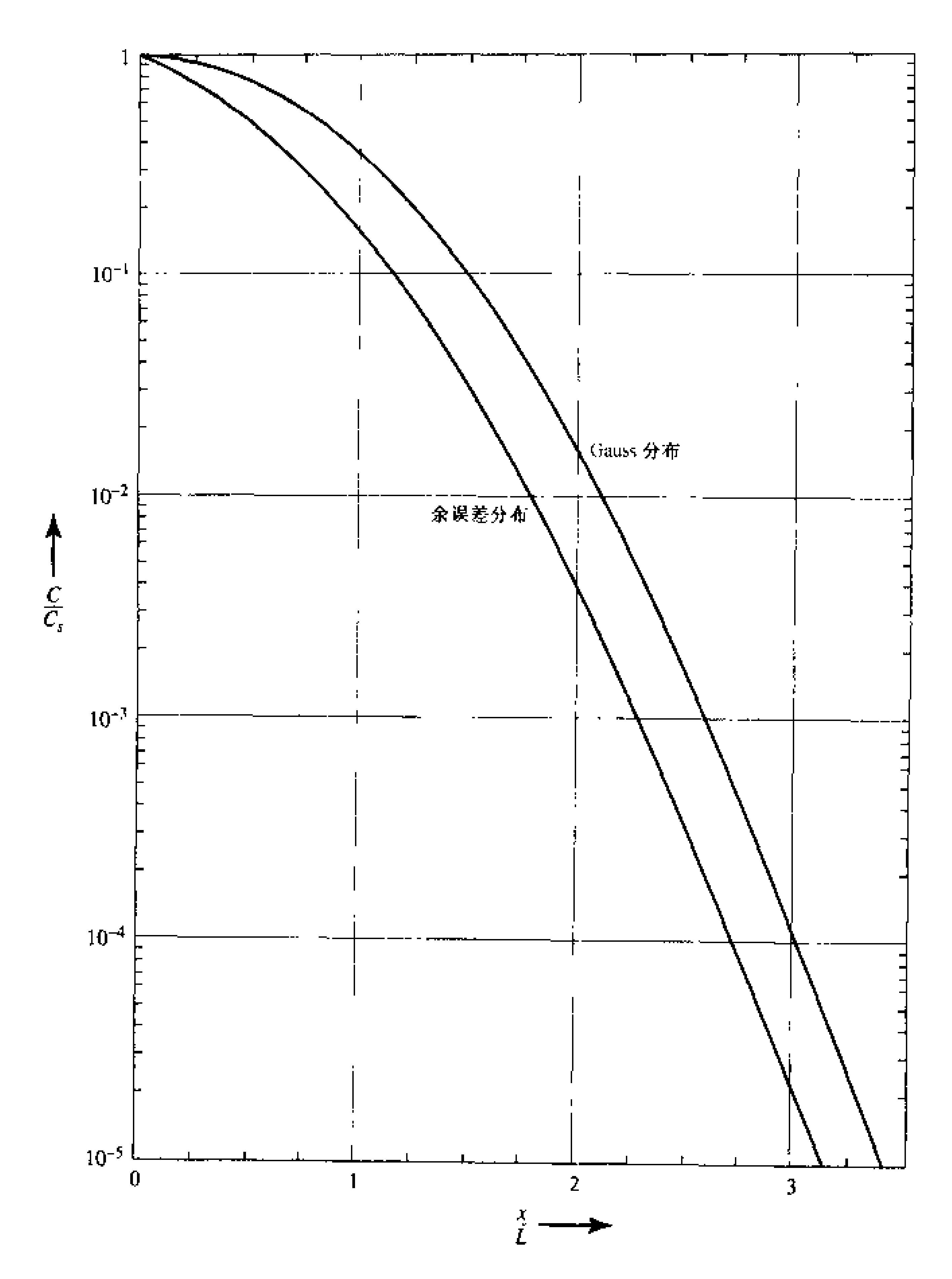


图 2.21 余误差分布和 Gauss 分布;垂直轴用峰值浓度 C_s 归一化,水平轴用特征长度 $L = 2\sqrt{Dt}$ 归一化

$$N' = \int_0^{\infty} C(x,t)\,\mathrm{d}x = 2\sqrt{Dt/\pi}\,C_s \tag{2.5.10}$$

等式(2.5.10)表明,与扩散深度一样,杂质的面密度随扩散时间 t 的平方根增加。

杂质再分布 杂质原子淀积后要经过随后的热处理进行再分布。

离子注入后,杂质原子具有 Gauss 分布(等式(2.5.1))。由于总的杂质原子数是固定的,随后的推进扩散使初始分布展宽,而峰值浓度下降。新的分布也是 Gauss 分布,但是它具有新的特征长度,与注入的特征长度 $\sqrt{2}\Delta R_p$ 和扩散的特征长度 $2\sqrt{Dt}$ 有关。将这些量以均方根的形

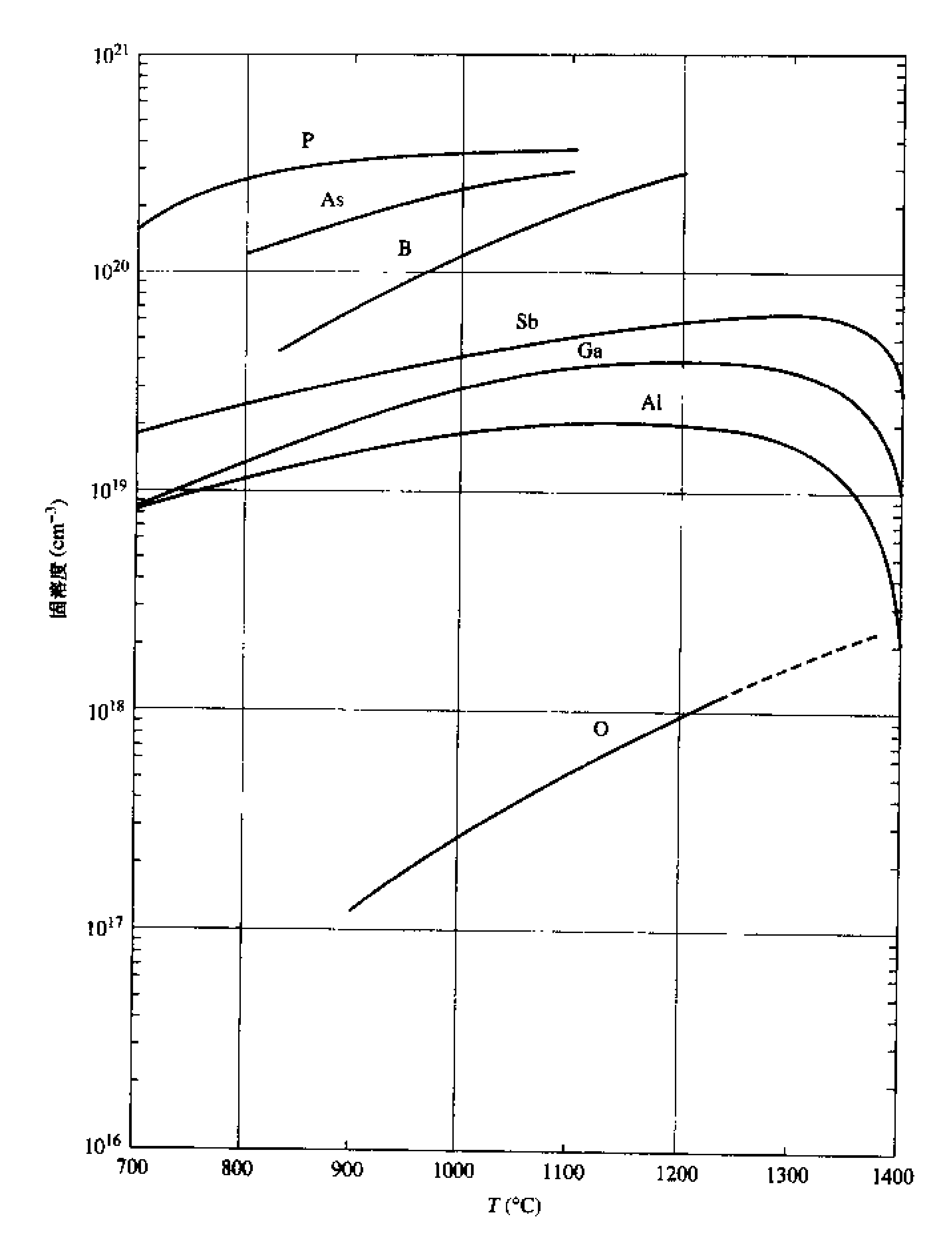


图 2.22 几种元素在硅中的固溶度与温度的关系[11]

式组合可得到新的特征长度 L'

$$L' = \sqrt{2\,\Delta R_p^2 + 4Dt} \tag{2.5.11}$$

用 L' 代替初始的注入特征长度 $\sqrt{2}\Delta R_p$ 就得到新的峰值浓度，于是最后的杂质分布为

$$C(x) = \frac{N'}{L'\sqrt{\pi}}\exp\left[-\left(\frac{x - R_p}{L'}\right)^2\right] \tag{2.5.12}$$

气相淀积的杂质原子再分布后也会形成 Gauss 分布。因为典型的推进扩散通常使杂质分布比气相淀积深得多，通常假定气相淀积在半导体表面形成了单位面积总浓度为 N' 的面杂质原子分布，将此近似作为推进扩散的初始分布，这样推进扩散只简单地对这一定量的杂质再分

布。对这样的边界条件，扩散方程的解也是 Gauss 分布

$$C(x,t)=\frac{N'}{\sqrt{\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right) \tag{2.5.13}$$

其中，特征扩散长度 $2\sqrt{Dt}$ 由推进扩散的时间和温度决定。

当杂质经过两次或更多次的扩散步骤连续再分布时，总的特征长度由每一步工艺的特征长度的均方根组合决定。例如对 m 步工艺

$$L'=\left(\sum_{i=1}^{m}L_i^2\right)^{1/2}=\left(\sum_{i=1}^{m}4D_it_i\right)^{1/2} \tag{2.5.14}$$

其中扩散系数取于每一步的扩散温度相对应的值(注意等式(2.5.14)中是对扩散长度的平方，而不是对扩散长度本身相加，最后是对每个 D_it_i 的乘积相加)。根据这种方式，在经历了一系列不同的热循环过程，比如说完成 IC 制造所需要的全部工艺后，就可以得到近似的杂质分布。

当掺入硅片的杂质原子的面密度一定时，采用气相预淀积得到的峰值浓度是离子注入得到的峰值浓度的两倍。这是因为对于气相预淀积而言，硅表面的杂质浓度最大，扩散只向硅片内部进行(假定阻止杂质从表面向外蒸发)。另一方面，离子注入的极大值浓度在硅表面下，因此扩散向表面和体内都在进行，使得固定数目的杂质原子分布在更大的范围内。

随着器件性能的提高，需要采用更浅的结。在某些情况下，离子注入工艺和随后消除晶格损伤的退火工艺形成的结深太大。为了获得非常浅的结，将一种具有高扩散系数的材料置于单晶硅上。将杂质原子注入到这层高扩散系数的材料中，然后杂质扩散进入下面的单晶硅中。因为上层材料的高扩散系数，杂质原子很容易扩散通过它，因此在硅表面的浓度几乎不变。另外，随着杂质从掺杂层材料扩散进入单晶硅，杂质很容易从掺杂层上层运动到界面，以补充扩散到单晶硅的杂质。这样掺杂层材料作为杂质源，保证了单晶硅表面的杂质浓度为常数，单晶硅中的杂质分布为余误差函数分布。尽管注入工艺会损伤掺杂层材料，但是对下面单晶硅的损伤则很小。

多晶硅掺杂源　一种可以很方便用做上述高扩散系数的掺杂层材料就是 2.6 节中将要讨论的多晶硅。多晶硅中晶粒间界的存在使得杂质原子比在单晶硅中的扩散快得多。高的杂质扩散系数及与硅工艺好的兼容性使得多晶硅近乎理想的杂质扩散源。

氧化硅掺杂源　另一种可供选择的掺杂方法是在硅片表面淀积一层含有掺杂杂质的氧化硅。然后，杂质原子从这一层玻璃态物质扩散进入硅中。形成掺杂氧化物的一种方便的方法是化学气相淀积(CVD)SiO_2，在淀积过程中将杂质原子加入氧化硅中。我们将在 2.6 节中更详细地讨论 CVD 工艺。也可以将杂质加入分散在有机溶剂中的玻璃粒子中，然后将这种材料旋转涂覆在硅片上，烘干，再把这层膜加热到约200℃，以去除残余的有机溶剂。用两种方法中的任意一种在硅片上淀积掺杂氧化硅后，必须采用推进扩散工艺使杂质经扩散进入硅中。

当使用掺杂氧化硅时，硅表面的杂质浓度在整个推进扩散过程中一般保持不变，其值等于氧化硅中掺杂浓度的固定比例。这一条件导致推进后再一次形成余误差函数分布(等式(2.5.8)和图 2.21)。

不同于多晶硅扩散源，杂质在玻璃态 SiO_2 中的扩散系数较低。但将硅表面处的杂质浓度

近似为常数是合理的,因为在氧化硅中杂质浓度非常高(达到百分之几十),只有非常小部分的杂质进入硅中,所以氧化物中的杂质浓度仍近似保持不变。

扩散系数的变化 因为杂质通过扩散物与点缺陷(主要为硅空位或硅填隙原子)之间的相互作用进行扩散,任何改变点缺陷密度的因素或改变与之相关的电荷分布的因素都可改变扩散过程。下面简单考虑两种重要的情况:(1)杂质的氧化增强扩散,(2)杂质的浓度相关扩散。

如果点缺陷浓度随晶体位置变化,扩散系数也与位置有关。在这种情况下,可把等式(2.5.5)和(2.5.6)结合起来将扩散方程表示为

$$\frac{\partial C}{\partial t}=\frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right) \tag{2.5.15}$$

而不是较简单的 Fick 定律的形式(等式(2.5.7))

等式(2.5.15)的解比前面讨论的简单的 Gauss 分布和余误差函数分布复杂。它的确切分布通常不能用简单的数学形式表示,需要用数值方法求解。但是,一些定性的说明可帮助了解这一复杂的扩散工艺。

晶体中靠近硅空位和填隙处存在局域态的电荷。这些点缺陷的带电状态与晶体中费米能级的位置有关,因而与杂质浓度和温度有关。因为扩散杂质原子与带电的、中性的以及点缺陷相互作用而进行运动,这种与 Fermi 能级的依赖关系影响到杂质的扩散系数。这时的扩散系数可写为杂质原子与不同带电状态点缺陷[14]相互作用的各分量之和。例如,与空位相互作用的扩散占主导地位时,有效扩散系数可写为

$$D_{\text{eff}}=h\left[D_i^0+D_i^-\left(\frac{n}{n_i}\right)+D_i^=\left(\frac{n}{n_i}\right)^2\right] \tag{2.5.16}$$

其中,h 是说明电场对扩散系数影响的参数,每一项与点缺陷不同的带电状态有关。

从等式(2.5.16)可看出,当表示带电空位作用的项与表示中性空位作用的项 D_i^0 可比拟时,扩散系数与杂质浓度有关。当 n/n_i 在扩散温度下接近于1时,不同 D_i 值的大小相当。例如,设扩散温度为1000℃;从图2.10得到的 $n_i\approx 9\times10^{18}\,\text{cm}^{-3}$。因此当杂质浓度高于这一值时,我们将观察到 D_{eff} 的增加。与浓度相关的扩散可使扩散系数增加10~20倍,使高掺杂区的扩散速率大大增强。那么,得到的杂质分布与适用于低掺杂浓度的 Gauss 或余误差函数不同。

在双极晶体管电路中经常可以观察到与杂质浓度相关的扩散现象,称为发射区推进效应。重掺磷或砷用于形成发射区,当与其相关的点缺陷引起基区杂质硼的扩散系数增加时,就会发生这一效应。这种基区杂质的发射区推进主要发生在发射区的正下方,但是也向两侧有小距离的推进。双极晶体管的设计及其工作原理将在第6章中详细介绍。图2.23是双极晶体管扫描电镜(SEM)截面图,在发射区的下面可以看到基区的发射区推进。类似于发射区推进造成的凸出的杂质分布在有些集成电路中的其他区域也会出现。

除了浓度相关的扩散,与其他机制导致的与点缺陷的相互作用也会引起扩散速率的增加。当硅氧化后,表面处有许多键断裂,形成点缺陷。点缺陷中的一部分向下面的硅内迁移,直至它们遇到另一种能与其湮灭的点缺陷。例如,空位和填隙原子复合(湮火),就像电子和空穴的复合一样。然而,在复合前,点缺陷可以在硅中运动相当长的距离,从而改变这一区域内任何杂质原子的扩散速率。这一效应造成了氧化增强扩散,当硅片上的部分区域被氧化,而另一

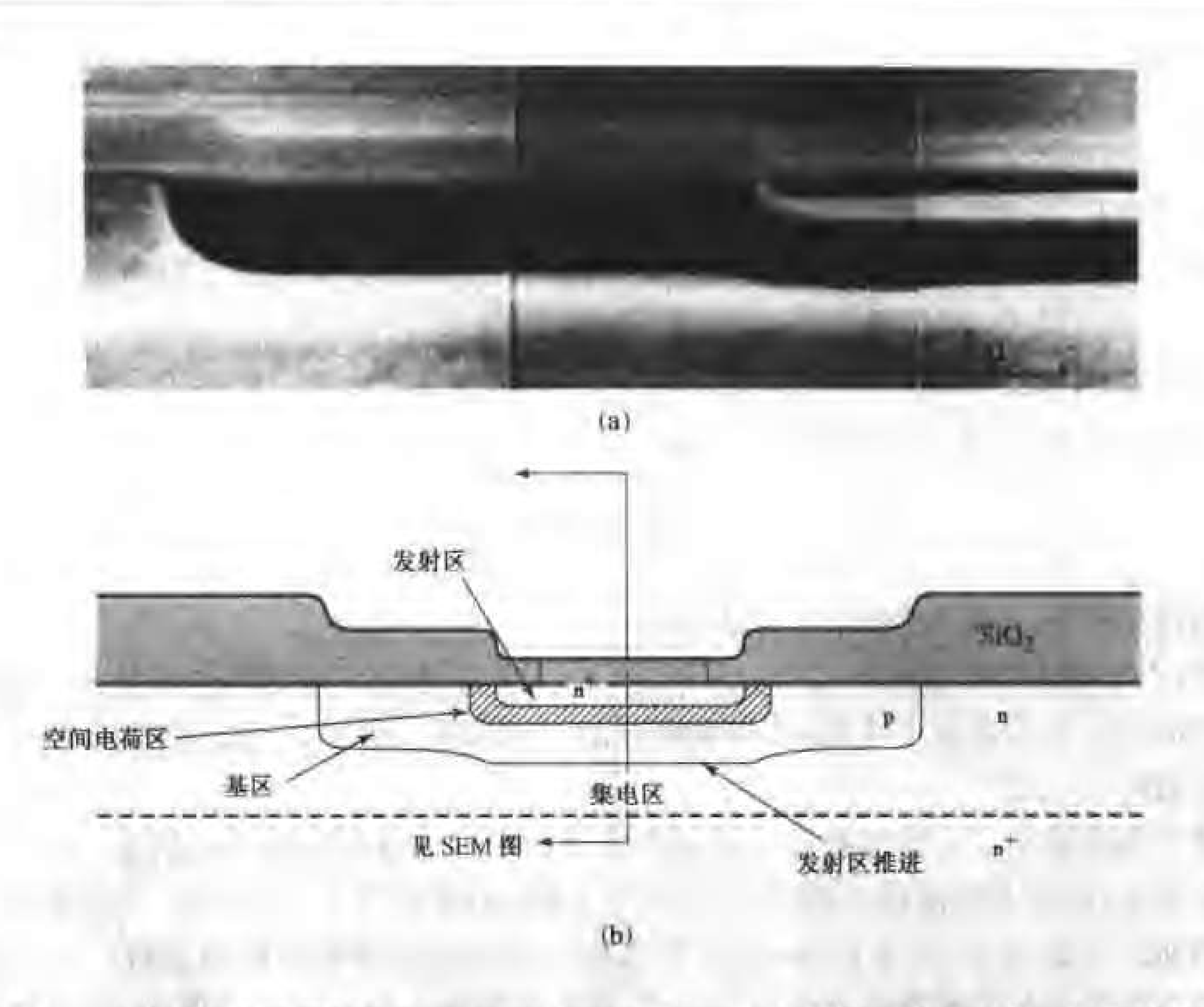


图2.23 (a)双极晶体管扫描电镜(SEM)的截面图照片;(b)各区域的示意图 掺硼的基区由于浓度相关的扩散效应向前推进(发射极推进),这一效应与发射区重掺磷有关[15]

部分区域被非氧化层覆盖时很容易观察到这一现象(图2.24)。硅片中氧化区域下面的扩散远大于未氧化区域下面(因而也没有高密度的点缺陷)的扩散。

除了氧化增强外,离子注入引起的晶格损伤也会大幅度增强扩散效应。在随后的热处理过程中,晶格损伤被消除,但是杂质原子也发生了扩散。在退火的初始阶段,晶格损伤促进了杂质的快速迁移(瞬态增强扩散),但是随着晶格损伤的消除,扩散减慢。将这些效应结合起来进行定量描述是非常复杂的,必须用计算机建模处理。晶格损伤与杂质运动间的相互作用在快速热退火工艺中尤其严重,快速热退火工艺(RTP)利用辐射效应只需几秒钟就能将硅片加热到高温,消除晶格损伤并激活杂质原子,同时使扩散降到最小。晶格损伤和点缺陷对杂质运动的影响在不同温度下变化很大。在进行杂质激活之前,适当消除晶格损伤(和点缺陷)可以将瞬态增强扩散降到最小。

固溶度 气相淀积过程中,进入硅表面的杂质数量受热力学过程的限制,最高只能达到其固溶度(图2.22)。但是,当用离子注入方法引入杂质原子时,杂质浓度可以超过其固溶度,因为这种情况不涉及热力学平衡的问题。但是,在随后消除损伤和激活杂质的退火过程中,杂质进入硅的位置的数量受到与退火温度相应的固溶度的限制。过剩的杂质可形成原子团或沉淀物。如图2.22所示,几种典型杂质原子的固溶度在常用温度范围内随温度升高而增加。因此,高温时杂质的固溶度会超过随后低温退火时的固溶度。当发生这种情况时,过剩的杂质趋于离开具有电活性的硅的位置。但是,如果温度足够低(例如室温),将过剩杂质从晶格中析出是非常慢的。实际上,如果杂质在高温下引入,硅片又很快冷却,掺杂浓度可以比室温下的固溶度高出很多。



图2.24 硅片的截面图，与未氧化的硅表面（顶部）相比，氧化后的硅表面（底部）内扩散的n型区（暗区）更深[16]

分凝 在热氧化过程中，硅被消耗，硅片中的任何杂质都必须在硅和生长的氧化硅中重新分布。在平衡态，$Si\text{-}SiO_2$界面两侧杂质浓度的比值是一个常数。用分凝系数（m）表示这个比值

$$m = \frac{C_{Si}}{C_{SiO_2}} \tag{2.5.17}$$

n型杂质磷和砷趋于向硅中扩散（$m>1$），被推进到不断生长的二氧化硅层的前面。而硼在硅表面区耗尽，进入氧化硅层中（$m<1$）。因为在热氧化过程中硅与二氧化硅的边界不断移动，界面处的平衡只有在氧化速率很低时才能达到。在典型的工艺条件下，分凝造成了与热平衡预计的杂质数值相差很大。杂质分凝也与杂质粒子在氧化物中的扩散速率有关。只有当杂质在氧化硅中扩散得很慢，在SiO_2气体界面处杂质的减少不能影响$Si\text{-}SiO_2$界面，平衡态才能达到。另外，还必须考虑通过扩散从体硅输运到界面处的杂质。界面处与体内的杂质浓度之比主要由氧化和扩散的相对速率决定；杂质原子在体内和界面间移动越慢，则在靠近硅表面处它的积累或耗尽越大。

这样，在确定表面分凝量时必须考虑三种重要的参数：(1)分凝系数m，(2)氧化速率与硅中杂质扩散系数的平方根（表示在硅中的扩散）之比，(3)杂质在硅中和SiO_2中的扩散系数之比。

快速热退火工艺和单片加工工艺 尽管IC工艺是基于一定数量硅片（通常是25片）的批量加工，一种称做单片加工工艺（即一次加工一个硅片）的相反趋势正在发展和实施。随着硅片直径的不断增加，为避免产生永久晶体损伤可容忍的热应力越来越小。在常规的退火炉中要求对硅片加热更慢一点；另外，与硅片和夹具相关的热容量在退火工艺中也会使热平衡延

迟。由于硅片达到稳态温度所需时间必须小于总的热循环工艺的时间,这些考虑限定了热循环工艺的时间下限。另一方面,在某一给定温度下允许的最长时间受到前面介绍的杂质原子的扩散的限制。采用低温,允许的处理时间长一些,但是某些工艺在高温下的效果会更好一些,例如,在高温下比低温下生长的氧化硅质量高。

为了用高温过程获得高质量的效果,并减小高温过程的时间,快速热退火工艺已经非常流行。在这种工艺中,低温下将硅片放入系统中,固定在热容量非常小的支架上(例如3个石英柱)。然后用灯来迅速升高硅片的温度,典型的加热速率为50~100℃/s。当然加热必须非常均匀以避免热应力造成晶体的损伤。硅片边缘的热耗散使硅片的均匀加热变得非常困难。但是,如果加热是均匀的,硅片可在10~20秒内被加热到相当高的温度(可能是1000℃)。退火可在几十秒时间内完成,这样能允许更高的退火温度而不造成杂质扩散。

尽管使用单片加工工艺似乎牺牲了批量加工的优势,但每个硅片短的加工时间又可补偿这种工艺一次只加工一片的缺点。例如,如果单片工艺可在1分钟内完成,它与100分钟(完成一批炉工艺的典型时间)内加工100片的批量工艺的产出率类似。另外,当加工少量特殊产品时,单片加工工艺尤其有优势,因为不需要那么多芯片填满炉子,限制了批量加工的经济效益。

即使在快速热退火不重要时单片工艺也有很大的优势。随着更先进的电路中的各层材料变得更薄,对层与层之间界面的控制变得很关键,在不同的工艺流程之间,将硅片保持在可控的、惰性环境下很有利。在单片加工中,不同的工艺反应腔围绕在中间的自动硅片处理装置(一种机器人)的周围。在真空或惰性气体环境下,硅片在这些工艺腔中传输。因为相关的工艺流程不需要将硅片暴露在空气中就可以完成,因此能对界面进行更好的控制。例如,在一个室中对硅片表面清洁处理后,在另一个腔中进行化学气相淀积或物理气相淀积。虽然这种复杂装置的可靠性还有问题,但在目前的IC加工中已经很普遍。

2.6 化学气相淀积

尽管集成电路的基本元件可以通过氧化、光刻和扩散形成,更复杂的结构要求在已形成集成电路的部分区域上面灵活地增加导电层、半导体层或绝缘层。采用淀积的绝缘层可用于避免杂质原子引入后的高温氧化过程,而淀积轻掺杂的单晶硅层或多晶硅膜可用于其他用途。这些材料可通过化学气相淀积(CVD)或物理气相淀积(PVD)的方法形成。在CVD工艺中,所有用于形成淀积层的反应物都以气态方式引入反应室中,硅片本身不参与反应。CVD膜的结构取决于衬底(非晶或单晶),也与淀积条件(主要是温度、淀积速率和气压)有关。CVD工艺可在很宽的温度范围内进行。通常采用加热衬底的方法来加速淀积反应,也可采用在淀积室中产生等离子体的方法来增加系统的能量。

2.6.1 外延

前面已经介绍可以通过离子注入或气相淀积和扩散的方法将杂质引入硅片。这些工艺可用来增加靠近表面层材料中的杂质浓度。然而,它们不能在表面层中形成低于内部的掺杂浓度。从理论上讲,采用1.1节中提到的补偿技术,通过掺入浓度近似相等、导电类型相反的杂质可以实现这个目的。然而,由于很难精确控制扩散过程,要获得杂质的均衡补偿几乎是不可

能的。采用均衡补偿方法的另一个缺点是在这一层材料中载流子的迁移率会下降，因为迁移率受总的电离杂质$N_d + N_a$散射的限制，而不是决定载流子浓度的净杂质数$|N_d - N_a|$的限制。

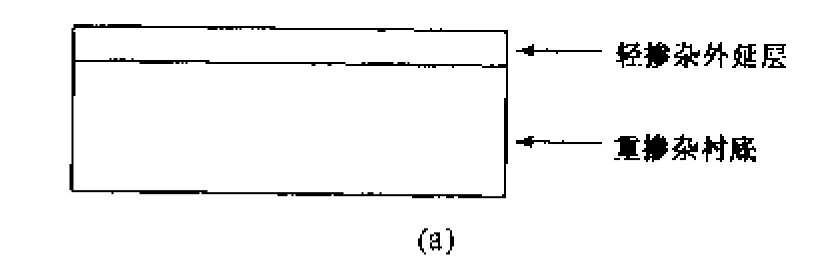
可通过外延工艺（图2.25）实现在重掺杂区上制备一层轻掺杂层，即在单晶硅片或衬底上控制生长单晶硅。为了生长外延层，将硅片放置在加热的反应室中，在其表面通入硅烷（SiH_4）或二氯氢硅（SiH_2Cl_2）气体。气体在硅片表面分解，在上面淀积出硅层。仅通过加热就能使硅烷分解

$$SiH_4(g) \rightarrow Si(s) + 2H_2(g) \tag{2.6.1}$$

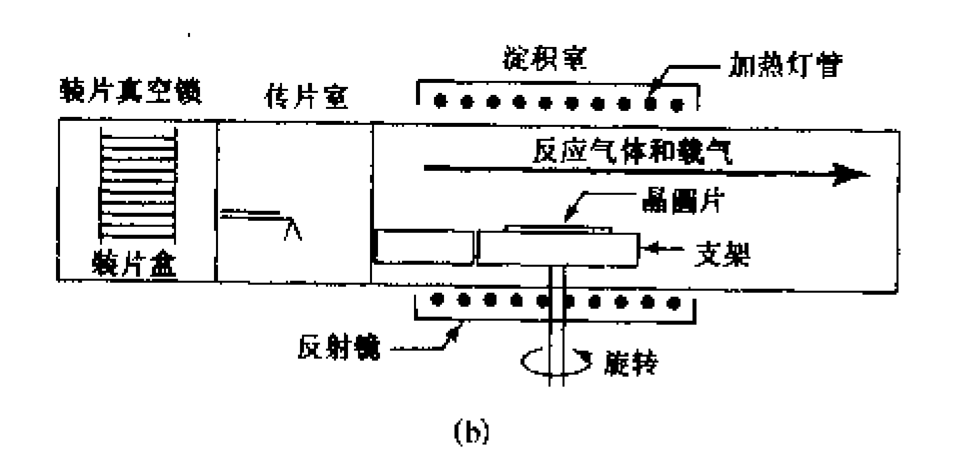
H_2是副产品，二氯氢硅分解会生成具有腐蚀性的副产品 HCl

$$SiH_2Cl_2(g) \rightarrow Si(s) + 2HCl(g) \tag{2.6.2}$$

注意等式(2.6.1)和(2.6.2)中的反应物都是气体，正好说明了化学气相淀积的含义，



(a)



(b)

图2.25 (a)轻掺杂外延层生长在重掺杂硅衬底上；(b)单片外延淀积系统，硅片放置在支架上被石英淀积腔外的红外灯加热

为了采用CVD方法在单晶硅上外延硅，必须将硅片加热到足够高的温度，以使得淀积的硅原子能获得足够的热能运动到低能的位置，在被后面到达的原子掩埋而不能移动之前，与衬底形成共价键并保持单晶硅的晶格结构。单晶硅生长或外延的温度通常在850℃至1200℃之间。外延层可比衬底的掺杂浓度低；因此，外延工艺提供了在高掺杂区上获得低掺杂浓度的一种方法。这种方法在优化双极晶体管结构时尤其重要。淀积时，可将含硅气体与含杂质的气体如砷烷（AsH_3），磷烷（PH_3）或硼烷（B_2H_6）等同时引入反应室，这样就可使杂质原子掺入生长的膜中。

2.6.2 非外延薄膜

除了外延硅，在IC应用中要用到许多种其他CVD膜。例如，在任何IC中，都需要有导电层来连接各器件。这些导电层必须与衬底隔离。铝是最常用的导电材料，但是它的溶点低

(660℃)且易于与其他元素反应,因此淀积铝后,应避免温度高于500℃的加热过程。如果采用一薄层硅作互连线,后面的热处理可以在1000℃或更高的温度下进行。许多IC应用中,这种耐高温的能力尤其重要。

多晶硅 淀积用做互连的硅与外延硅所用的方法类似。然而,因为这些硅层不是直接淀积在单晶硅上,而通常是在非晶SiO_2上,所以它们不能外延生长。这些CVD膜由许多小晶粒组成(通常直径为50~100nm),因此称为多晶硅。

多晶硅在MOS工艺中具有特殊的重要性,它能在硅栅MOSFET中用做晶体管的电极,这将在第9章中介绍。除此之外,多晶硅还可在双极工艺中用做扩散源(见6.5节)。在这一应用中,多晶硅是在单晶衬底上用大约600℃~700℃低温CVD方法形成的。在这一温度下,由于硅表面有残余氧化物,且硅原子在表面迁移所能获得的热能有限,因此生长的是多晶硅而不是单晶硅。在随后的工艺中,这层多晶硅的性质与该多晶硅和下面的单晶硅衬底间的薄氧化层(约1个原子层厚)有密切关系。

非晶硅 当CVD硅在更低的温度(低于600℃)下淀积时,不管衬底如何,形成的都是非晶薄膜。非晶硅结构短程有序(一般只有几个原子间距),其中不含结晶相。在中等温度(≥600℃)的热处理过程中,非晶态硅晶化形成多晶硅。淀积的非晶硅薄膜再结晶,这一过程在许多应用中变得越来越重要,例如在液晶显示中的有源晶体管阵列。

在600℃以下可以通过硅烷热分解来形成非晶硅。然而,在相当低的温度(300℃)范围内,通过等离子的能量也可分解硅烷形成非晶硅膜。在这么低的温度下分解硅烷会导致相当多的氢存在于淀积的薄膜中,这对薄膜的电学特性有很大的影响。

绝缘薄膜 绝缘薄膜和导电薄膜一样可用CVD方法制备。IC工艺中特别重要的CVD薄膜是二氧化硅(SiO_2)和氮化硅(Si_3N_4),硅烷(SiH_4)或二氯氢硅(SiH_2Cl_2)与氧气或笑气(N_2O)反应可制备SiO_2,与氨气(NH_3)反应可制备Si_3N_4。尽管CVD方法制备的氧化硅不如热生长的氧化硅纯度高,电学特性也要差一些,但它不需要高温过程。例如,在淀积铝或铜之前,CVD氧化硅可淀积在多晶硅互连线层上,这样能避免热氧化时杂质的扩散。CVD氧化硅也可在更低的温度下淀积在铝或铜上,以隔离连接有源区器件的不同的导电层,另外,还可用来保护制作完毕的集成电路不受外界污染。在后一种情况下,氧化硅中通常要掺磷,以阻止任何污染通过氧化层进入到电路中。通常在约400℃的淀积温度下,通过硅烷与氧反应制备用于导电层间隔离的氧化硅。反应过程为

$$SiH_4(g) + O_2(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{2.6.3}$$

与热氧化硅相比,淀积的氧化硅不够致密,化学惰性也不够好,但是在需要避免高温工艺时非常有用。

集成电路器件间的隔离要用到局部氧化工艺(LOCOS),因为氮化硅不像硅那么容易氧化,通常用它来限制热氧化的区域[17]。在LOCOS工艺中,氮化硅层淀积在硅衬底上,光刻后在制作器件的区域留下氮化硅(图2.26(a))。在需要生长厚氧化硅以用于器件隔离的区域中除去氮化硅。通常在氮化硅和硅片之间加一薄层SiO_2可以减小氮化硅层的应力,避免在硅片上产生缺陷。当氮化硅图形化后,将硅片插入到氧化炉中,裸露硅的区域通常在水蒸气或高温的(H_2:O_2)气氛中生长出一层厚的氧化硅(图2.26(b)),氮化硅阻止了器件区域的氧化。而在裸露的硅上可以生长出超过一个微米厚的SiO_2层,大约只有几十纳米的氮化硅转变成

SiO_2。氧化后，去掉氮化硅上的薄氧化层，然后腐蚀掉氮化硅和薄的应力缓冲氧化层，露出制作器件的硅区域(图 2.26(c))。图 2.26(d)给出了用 LOCOS 工艺形成的器件区的扫描电镜图。氧化层厚度从隔离区逐渐减小过渡到有源区。这一锥形过渡区在制作连续薄膜时避免了锐利的边界。然而，对尺寸很小的器件，这种锥形区域是有害的，因为这限制了可形成非常小器件的尺寸(亚微米尺寸)。

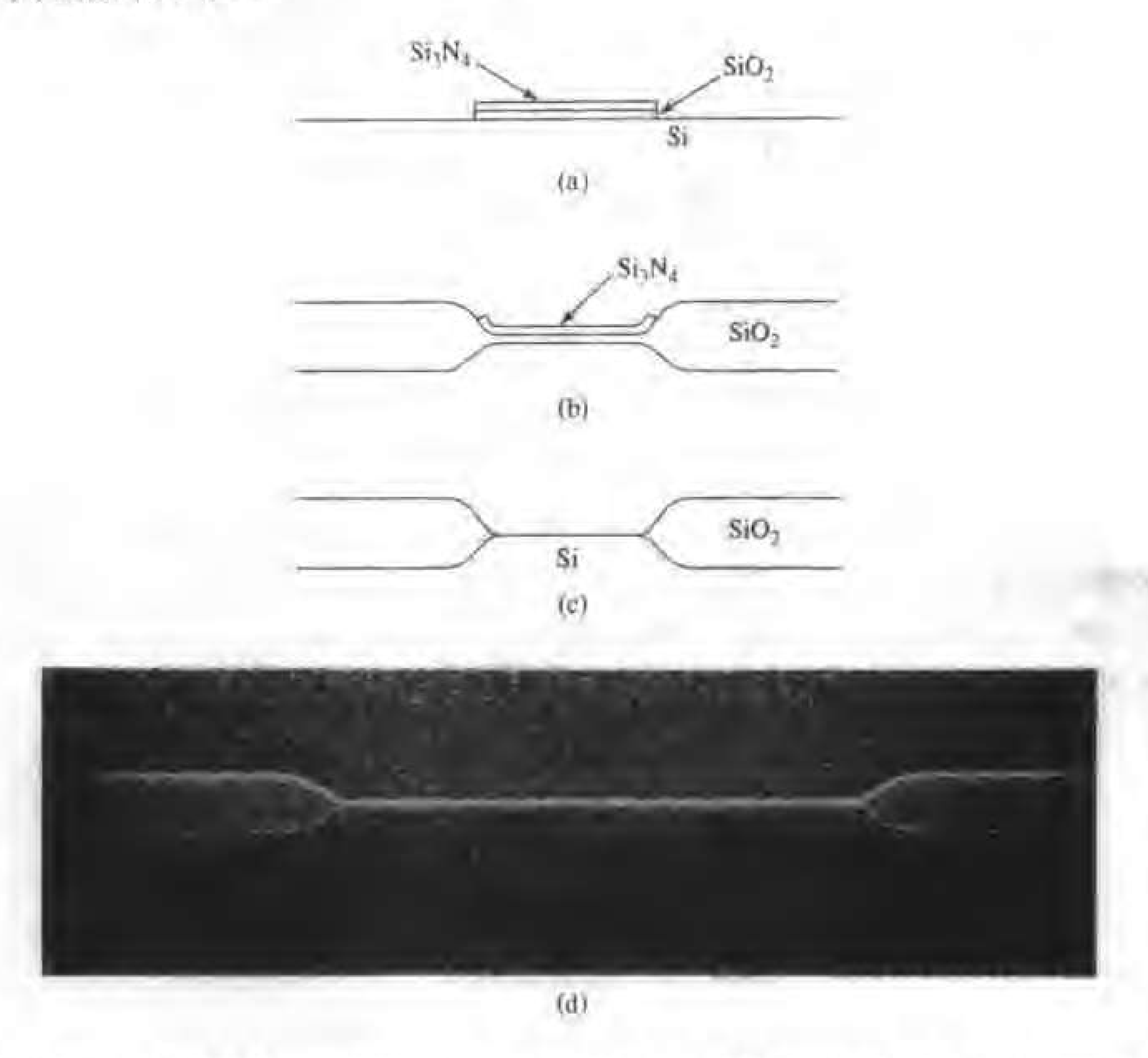


图 2.26 硅的局部氧化(LOCOS)工艺。(a)确定不需要氧化的区域，由 Si_3N_4 和应力缓冲氧化硅覆盖；(b)在裸硅区生长厚的氧化硅层；(c)去除应力缓冲氧化硅和 Si_3N_4 以制备器件；(d)采用 LOCOS 工艺加工(b)工序时的扫描电子显微镜(5000×)照片

为了制作适用于图形非常精细的隔离区，可采用沟槽隔离的方法，如图 2.27 所示。所谓沟槽隔离，就是在被隔离器件周围刻蚀出非常窄且相当深的沟槽。一般采用定向反应离子刻蚀形成几乎垂直的沟槽侧壁，使得沟槽上端能够得到最小的尺寸，而且沟槽也足够深(图 2.27(b))。当在侧壁和沟槽的底部形成薄的高质量绝缘氧化物(图 2.27(c))后，沟槽内或者被氧化硅或者被多晶硅填满。对深的沟槽，通常用多晶硅，因为它的热膨胀系数几乎与硅相同，这样可减小热应力和硅片的弯曲。对浅的沟槽，利用氧化硅时对应力的影响最小。当采用氧化硅时，通常采用高密度等离子体淀积(图 2.27(d))。高密度等离子体良好的方向性，可以填满非常深且窄的沟槽。在淀积过程中，离子轰击破坏了表面淀积层上的弱键，提高了氧化硅的质量和密度。

反应动力学 CVD 工艺通常是在开管的反应室中进行的，如图 2.25(b)所示。气流连续通过反应室，气相的反应产物与未反应的气体一起被抽走。通常利用载气推动反应气体流过

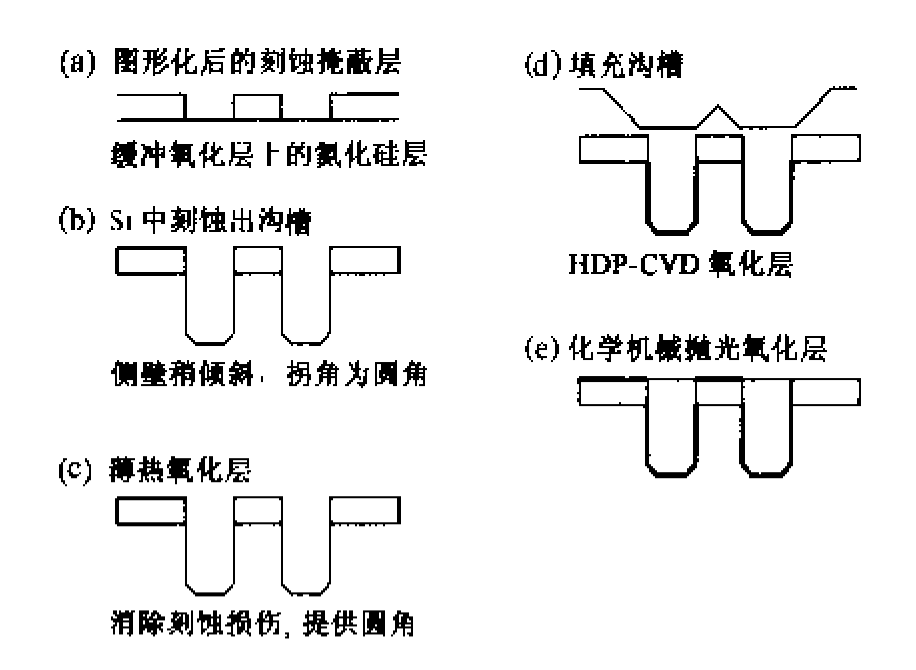


图 2.27 沟槽隔离用于相邻器件间距很小时的情况，当掩蔽材料图形化后(a)，用定向反应离子刻蚀工艺刻蚀沟槽(b)。形成一层薄的高质量氧化硅(c)，用多晶硅或氧化硅填满沟槽((d)中所示为氧化硅)。多余的材料用化学机械抛光去除(e)

反应室。只要反应气体在低温下不发生反应，在进入反应室前先将它们混合。在硅片表面，反应气体必须扩散通过一气相边界层(图 2.28(a))。这一边界层是位于自由流动区和反应室壁及夹具之间的过渡区域，该处气体的流速受到粘滞力作用而下降，粘滞力阻碍了气体向前流动(图 2.28(b))。

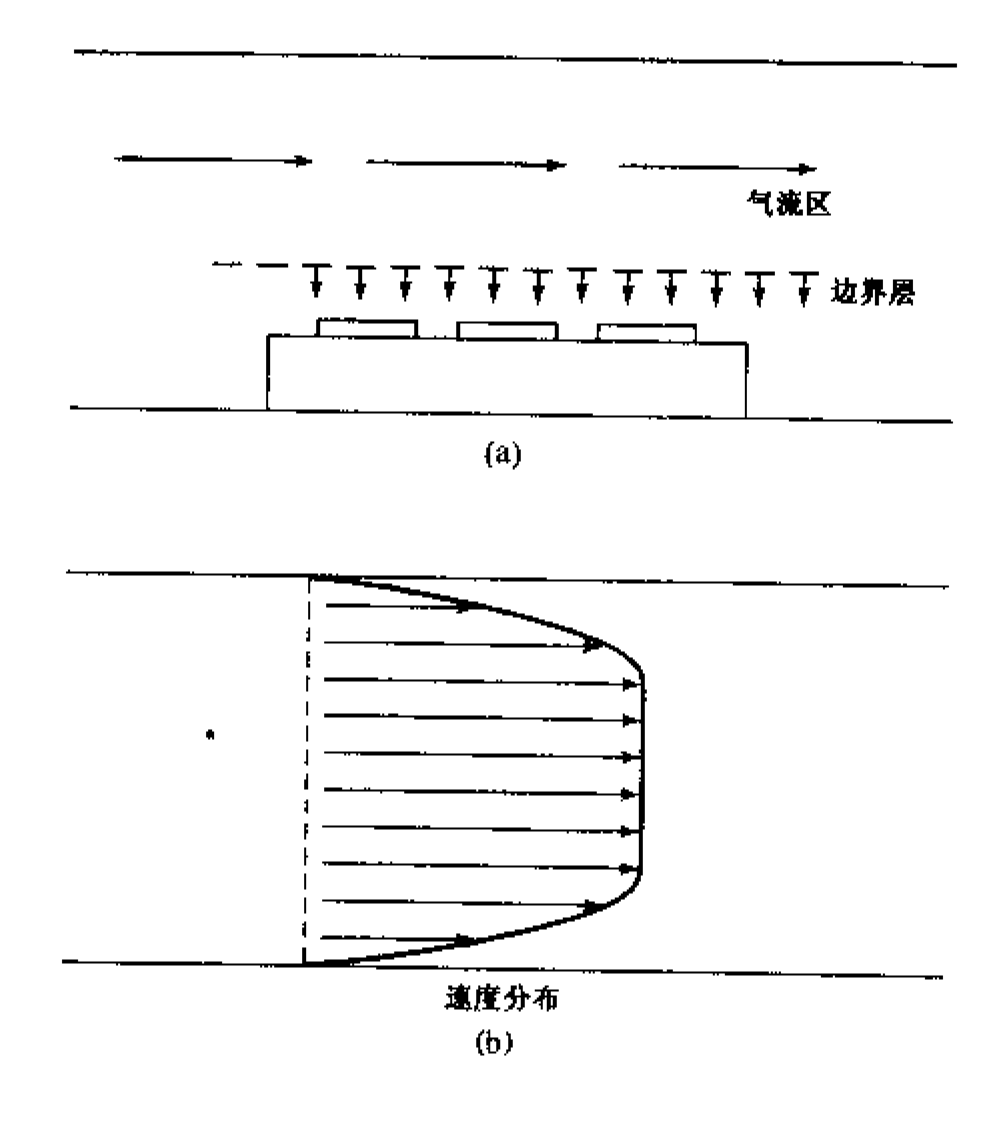


图 2.28 (a)沿水平方向的开管反应室截面图，表明气流平行于硅片表面，而在边界层位置气流几乎与硅片表面垂直；(b)通过反应室的气流速率分布

反应气体必须通过边界层才能到达硅片表面，加热可促进反应的进行。气体达到的所有加热的表面都会发生淀积反应。如果反应室壁是热的，则与硅片表面一样会有薄膜淀积，这是不需要的，而且会带来麻烦。无论通过边界层的扩散速率还是表面的反应速率都会限制总的淀积速率。在2.3节中热氧化的情况下，我们讨论过类似的传输速率限制。不同于热氧化，CVD的扩散过程发生在气相层（即边界层）而不是固相的SiO_2层上。氧化时可写出通过边界层扩散的分子流量[$F(1)$]和通过表面反应的分子流量[$F(2)$]（图2.29）

$$F(1) = D\frac{C_g - C_s}{\delta} \tag{2.6.4}$$

和

$$F(2) = k_s C_s \tag{2.6.5}$$

在等式（2.6.4）中，D为反应气体的气相扩散系数（与温度有弱的依赖关系），δ为边界层厚度，C_g和C_s分别表示气相边界层的外边界和固体表面附近的浓度。在等式（2.6.5）中，k_s为表面反应速率常数，与温度的倒数成指数关系，激活能为E_a[$k_s = k_{s0}\exp(-E_a/kT)$]。

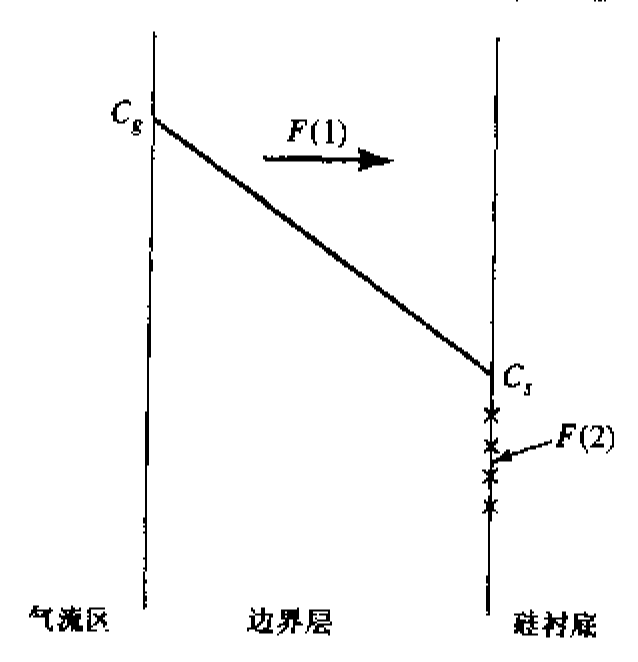


图2.29 反应气体以速率$F(1)$扩散通过边界层进入到硅片表面，并在那里以速率$F(2)$反应

稳态时，$F(1) = F(2) = F$，总的淀积速率R_d可写为

$$R_d = \frac{F}{N} = \frac{C_g/N}{\delta/D + 1/k_s} \tag{2.6.6}$$

N为淀积的薄膜中单位体积内的原子数。等式（2.6.6）分母中的第一项表示气相扩散的阻力，第二项为表面反应的阻力。受到表面限制的反应强烈地依赖于k_s，在低温和低压（$p \ll 1\text{atm}$）的CVD工艺中占主导地位。然而，k_s随温度增加很快，在高温下表面反应不再限制总的反应过程；通过边界层的气相扩散起到限制作用。典型的温度关系示于图2.30。当压力增加时（达到一个大气压）气相扩散变得困难。因此，高压容易导致气相扩散限制总的反应过程。

CVD反应室的几何形状也会影响到哪一个限制过程起主要作用。在接近一个大气压下，反应室气流几乎平行于硅片表面时，如图2.25（b）所示，气流很容易控制，而温度很难控制。因此，通常需要设计对温度不敏感的气相扩散过程作为总的CVD反应限制过程。在水平反应

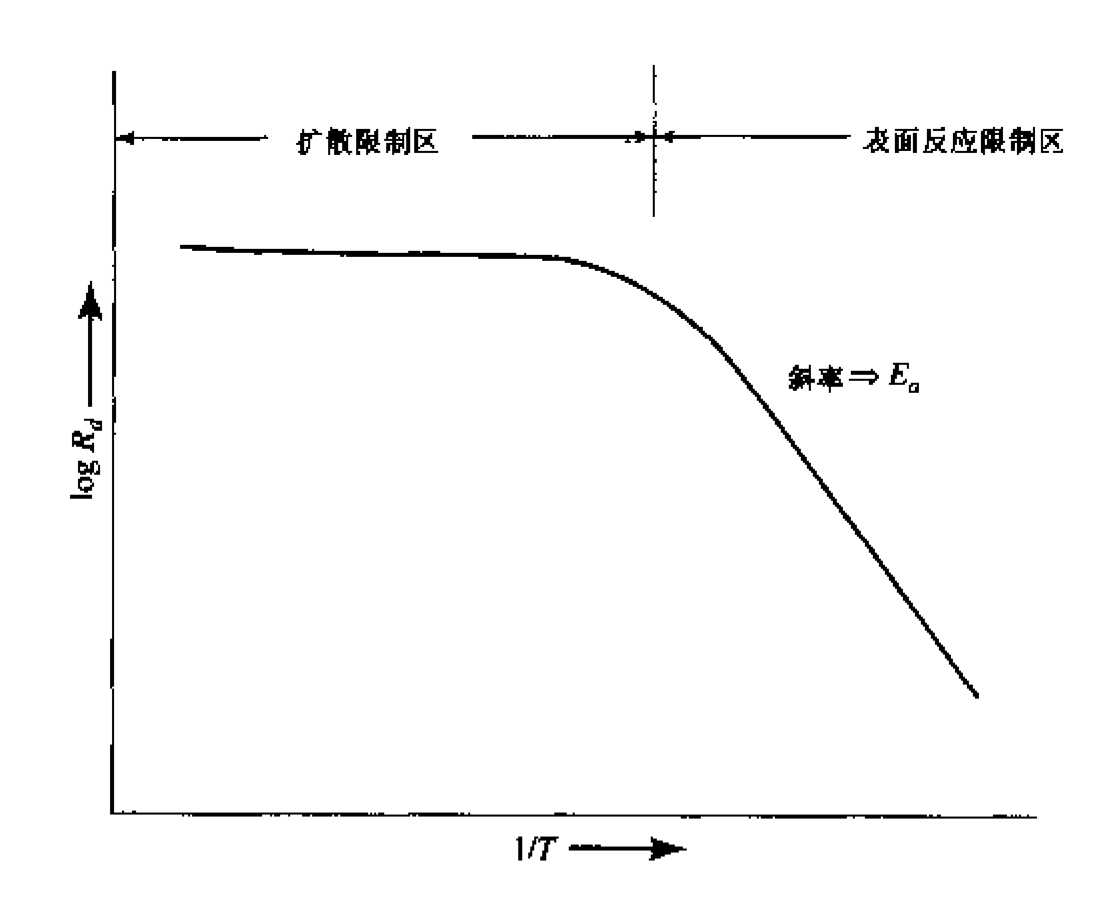


图2.30 总的淀积速率R_d与温度的倒数之间典型的关系曲线;低温下,表面反应速率限制了淀积速率,而在高温下,气相扩散(质量转移)限制了反应速率

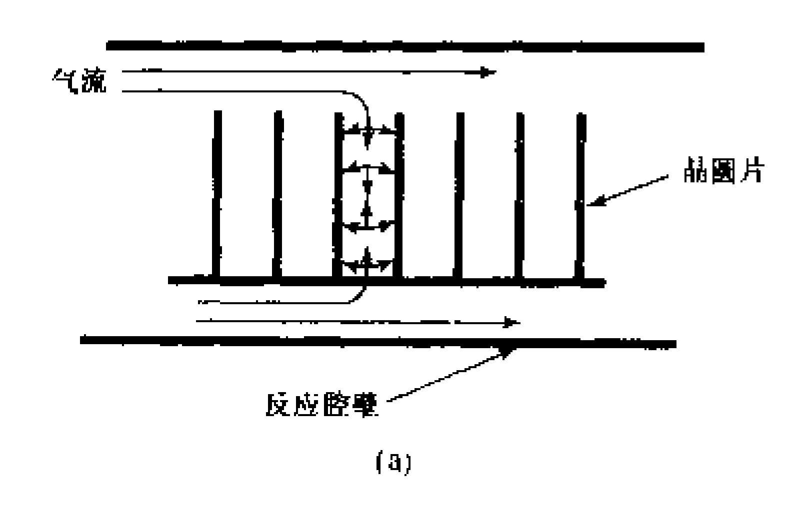
室中,产出率十分有限,尤其是一次只能加工一个硅片的单硅片反应室。

在可容纳许多硅片的反应室中,硅片就像在扩散或氧化炉中那样紧密地排列(图2.5)。在这种几何结构下,气体流动很难控制,因为气体首先要流过围绕硅片的环形空间,然后进入硅片间的窄缝隙中(图2.31(a))。因此扩散速率限制了反应,导致薄膜厚度不容易均匀。另一方面,这种炉子的温度能够控制得非常好,所以可工作在表面反应限制区域。因此,硅片紧密排列的高容量反应室一般在较低的温度下工作,表面反应速率限制了淀积速率。如果工作在低压下(≈0.133kPa),扩散速率限制过程会进一步减小。因为气相扩散系数与压力的倒数成正比,等式(2.6.6)式中分母的第一项变得更不重要。

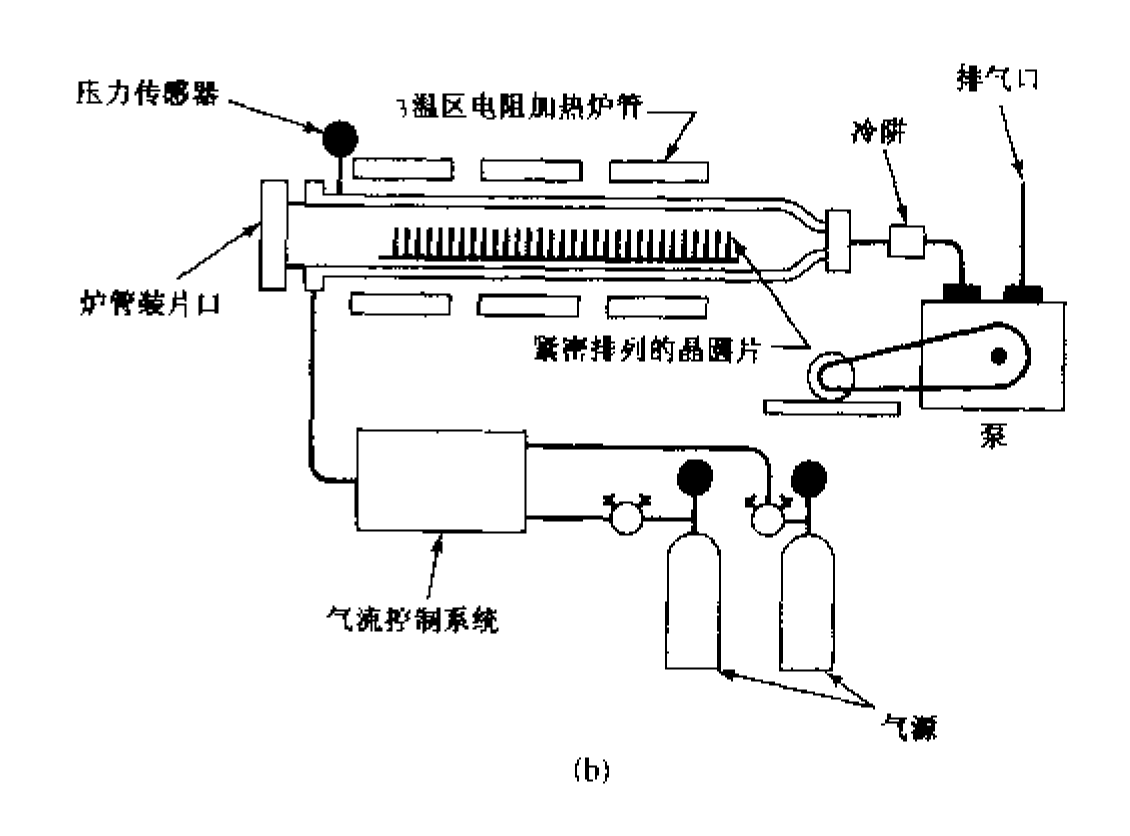
高容量低压CVD(LPCVD)反应室一般用于淀积多晶硅和氮化硅。LPCVD系统的基本结构如图2.31(b)所示。反应室按图2.31(b)所示方向放置,气流主要在水平方向,硅片垂直放置。另一种方式是将反应室旋转90°,气流主要在垂直方向,而硅片水平放置。后者更容易实现自动硅片处理。

这种类型的反应室在硅外延中还没有广泛应用,因为为保证单晶硅生长一般需要更高的温度。正如前面已经提到的,采用高温与避免扩散限制反应之间有矛盾。然而,这种结构的超高纯度的反应室可用做低温度下的专门的外延淀积。

等离子增强CVD 对某些应用,必须在低温下淀积薄膜。例如,芯片中铝或铜互连线上的钝化层必须在低于400℃的温度下淀积。在这一低温下,热淀积或者速率太低,或者薄膜不能很好地覆盖住不规则的图形。在这些情况下,仅靠热分解的CVD是不够的,必须给硅片表面提供额外的能量使得化学反应能够进行,或者改善被淀积材料的特性。高频电场可以给气体的混合物提供电能以产生等离子体,在等离子增强CVD(PECVD)中加强了淀积过程,与2.4节中讨论的增强刻蚀过程一样。在这两种情况下,离子轰击非常重要。在淀积过程中,离子轰击可改变淀积层的性质,甚至可以对淀积的薄膜进行溅射。



(a)



(b)

图2.31 (a)高容量反应室的气体流过反应室壁和硅片间的环形空间，然后在紧密排列的硅片间扩散；(b)LPCVD反应室的基本结构

等离子增强CVD在IC制造中越来越重要，可用来淀积氧化硅、氮化硅和硅。因为PECVD是低温工艺，因此薄膜的性质与高温热氧化薄膜的性质不同。PECVD方法制备的氧化硅和氮化硅通常不满足化学计量比(即不是SiO_2或Si_3N_4)，硅的结构也不好描述。

等离子的特性与许多独立的变量有关，例如电子浓度、电子能量分布、气体密度和等离子体内激发态的时间。这些微观变量由宏观参量来控制，如(1)反应室的几何形状，(2)用来激发等离子体的高频电源的频率和功率，(3)泵的抽速，(4)电极温度，(5)反应气体和稀释气体的流速。很不幸，宏观与微观参量之间并没有直接的定量关系。

在最简单的等离子体反应室中，等离子是靠激发所有反应气体和稀释气体的混合体产生的。激发态的中性原子团扩散到硅片表面，离子被等离子体和硅片间的电场加速向表面运动。而更先进的反应室中，人们尝试分别控制等离子体的产生、化学反应和离子轰击。在这些远程的等离子反应室中(图2.32(a))，等离子体在一个区域产生，通常激发气体本身不参与淀积反应(例如氧)。然后将等离子体传输到硅片表面。在靠近硅片表面处，它与更多的反应气体混合(如硅烷)，并在硅片表面形成需要的薄膜(例如SiO_2)。等离子体的这种远程产生方式更利于控制反应过程，灵活性也更大。它减少了不必要的反应(和颗粒)，更重要的是它可以控制

表面的离子轰击。

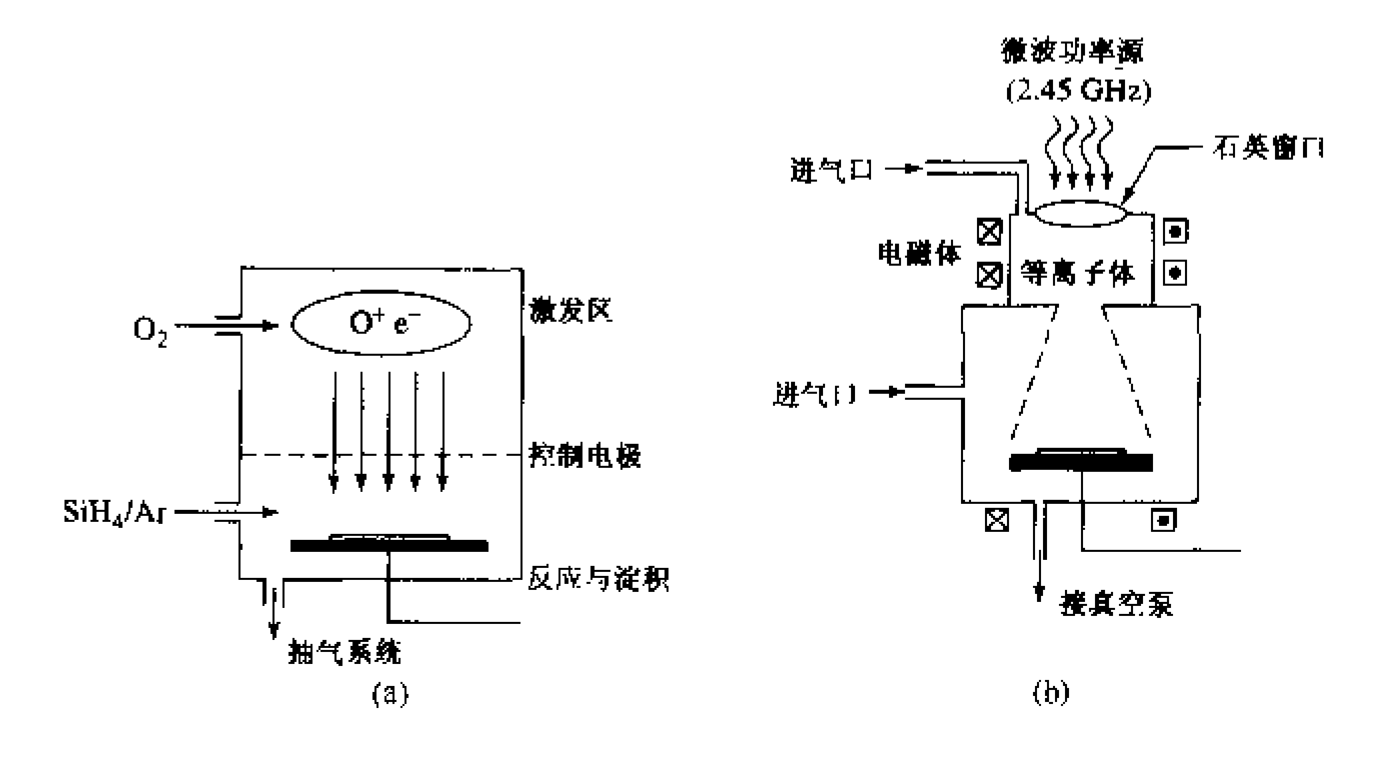


图 2.32 (a)远程等离子增强 CVD 反应室的截面示意图,等离子体的产生、化学反应和离子轰击被部分分开;(b)电子回旋共振、高密度等离子反应室的截面示意图

当离子的平均自由程接近反应室的特征尺寸时,离子几乎不受到散射,大部分能沿着电场的方向接近表面。它们可以从淀积的薄膜上轰击出成弱键的原子,提高了薄膜的密度和质量,高能离子能将阻止继续生长的表面淀积物溅射掉。包含足够多的离子以至能明显改变淀积薄膜性质的等离子体称为高密度等离子体(HDP)。图 3.32(b)给出了一种 HDP 反应室的截面图 在这种电子回旋共振(ECR)反应室中,等离子体与电场和磁场相互作用约束了大部分等离子体,实现了有效的能量吸收和等离子体的产生。

2.7 互连和封装

2.7.1 互连

为了制造集成电路,用平面工艺制作的单个器件必须用导电线相互连接起来,如图 2.33 所示。这一过程通常称为互连或金属化。随着单个晶体管的性能的提高,总的电路性能会受到晶体管之间的互连的限制,甚至超过了晶体管本身的限制。

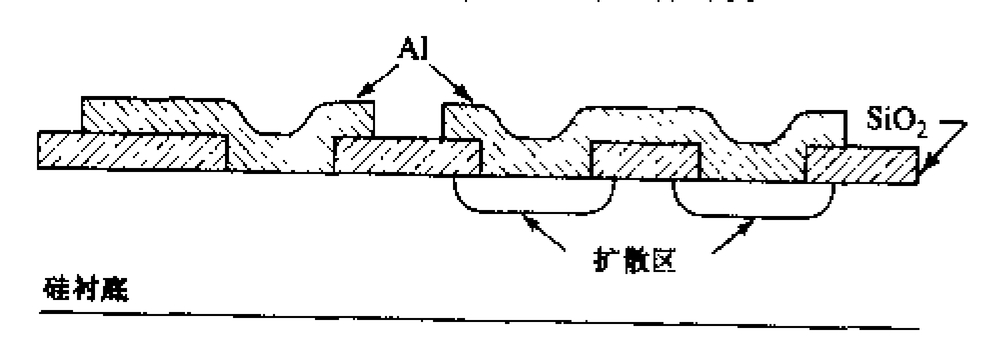


图 2.33 一薄层铝可用来连接半导体器件的各种掺杂区域

最简单且应用最广泛的互连方法是减法工艺。首先,去除接触孔处的 SiO_2 层以暴露出硅。然后,通常采用物理气相淀积(PVD)在表面淀积一层金属。被淀积金属的固态源在真空室中,或者用电子束轰击蒸发或者在低压气氛中用离子束轰击(溅射)。气态的金属原子运动

到硅片表面,在那里凝结形成均匀的薄膜。最常用的金属是铝或铝的合金(含有百分之几的硅或铜),不需要的区域用前面讨论过的光刻和刻蚀去除。通常采用的是各向异性干法刻蚀技术对铝进行刻蚀,当特征尺寸较大时,也可在磷酸溶液中用湿法腐蚀。

对于浅结,即使金属向硅中略微渗透也是不能容忍的。硅在铝中的溶解度较小,所以衬底中少量的硅溶解在纯铝中就能达到其固溶度。在衬底中这样形成的空洞会被铝填满。当这一"铝穿刺"产生时,铝可以穿透 pn 结而产生大的漏电通道。过去,在淀积铝膜时,有意在其中加入百分之几的硅以阻止衬底硅在铝中的溶解。然而对非常精细的图形,采用掺硅的铝就不太有效,因为超出固溶度的那部分硅会在非常小的接触区沉淀下来,占据其中相当大的一部分面积,从而明显增加了接触电阻。因为这个原因,通常要在铝金属化层和硅之间增加一层电极。正如下面要讨论的,通常采用硅化物接触以降低扩散区的串联电阻。其他方法包括硅和铝之间加一层薄的导电势垒层。最常用的是钛和钨的合金或钛与氮的合金。

当结变得更浅时,扩散区的电阻率增加,流过器件的电流受到与器件有源区串联的扩散区横向电阻的限制,而不是受到有源区本身的限制。为了减小这一串联电阻,可在扩散区顶部(平行地)放置一层低阻材料。其中一种方法是在裸露的硅上形成自对准金属硅化物,如图2.34(a)所示。将金属层如钛,淀积在整个硅片上。然后对硅片加热,使得钛与裸露的硅发生反应形成 $TiSi_2$,而在氧化硅上的钛则不发生反应。在只腐蚀金属而保留硅化物的腐蚀液中,将没有参与反应的金属从氧化硅上湿法去除。进一步的热处理将硅化物转变为低阻态,因此低阻的硅化物($\rho \approx 15\mu\Omega \cdot cm$)保留在硅上,而绝缘的氧化硅上没有导电材料。采用自对准硅化物的形成工艺,不需要额外的面积用于自对准光刻掩膜,可形成非常紧密的排列。其他金属,如钴,也可用于形成自对准硅化物。在 MOS 晶体管的电极上形成低阻层时,硅化物工艺特别有用。图2.34(b)所示为单晶硅源区、漏区及在多晶硅栅电极上形成的硅化物。

在几乎所有的 MOS 电路和许多双极电路中,通常一层互连由多晶硅组成,另一层由铝或铜组成。这两层互连线再加上硅片表面可能的扩散互连线,在垂直方向上电流可在三个不同的层中流动——这是成千上万个器件组成的电路中的一个重要自由度。我们将在下面看到,在复杂的电路中还要用到其他的金属层。

但是,多晶硅的电阻率大约在 $500\mu\Omega \cdot cm$ 左右,所以在长的多晶硅导体中会有相当大的电压降。然而更重要的是,与长多晶硅互连线有关的 *RC* 时间常数以及它与衬底之间的电容会使 IC 中信号传输速度减慢。因此,多晶硅只能用于非常短的局部互连。比多晶硅导电性能更好的其他材料用于长的互连。难熔的金属硅化物,如硅化钨(WSi_2),硅化钽($TaSi_2$)和硅化钛($TiSi_2$)及难熔金属本身,尤其是存在高温工艺时,通常被用做中等长度的互连线。

多层互连 随着现代集成电路复杂性的日益提高,晶体管和其他电子器件的互连变得更困难了。尤其是在逻辑电路中,信号必须经常从芯片的一部分传输到另一部分。当需要长距离传输信号时,必须采用金属互连,因为金属的电阻率低。然而,因为有许多交叉互连,一层金属是不够的,必须采用多层互连。一般来讲,第一层金属用于连接器件本身(通常是在硅或多晶硅的势垒层上)。第一层金属上通常覆盖有二氧化硅绝缘层。去除金属层之间互连区上的绝缘层,然后第二层金属淀积在上面并图形化。然后再淀积绝缘层和金属,再进行图形化,这一过程不断重复,最终构成复杂的多层金属化互连系统。具有五层金属的电路已经很普遍,人们希望采用八层或更多的金属层。

图2.35 给出了三层互连系统的透射电镜截面图。看起来互连系统占据了芯片的大部分

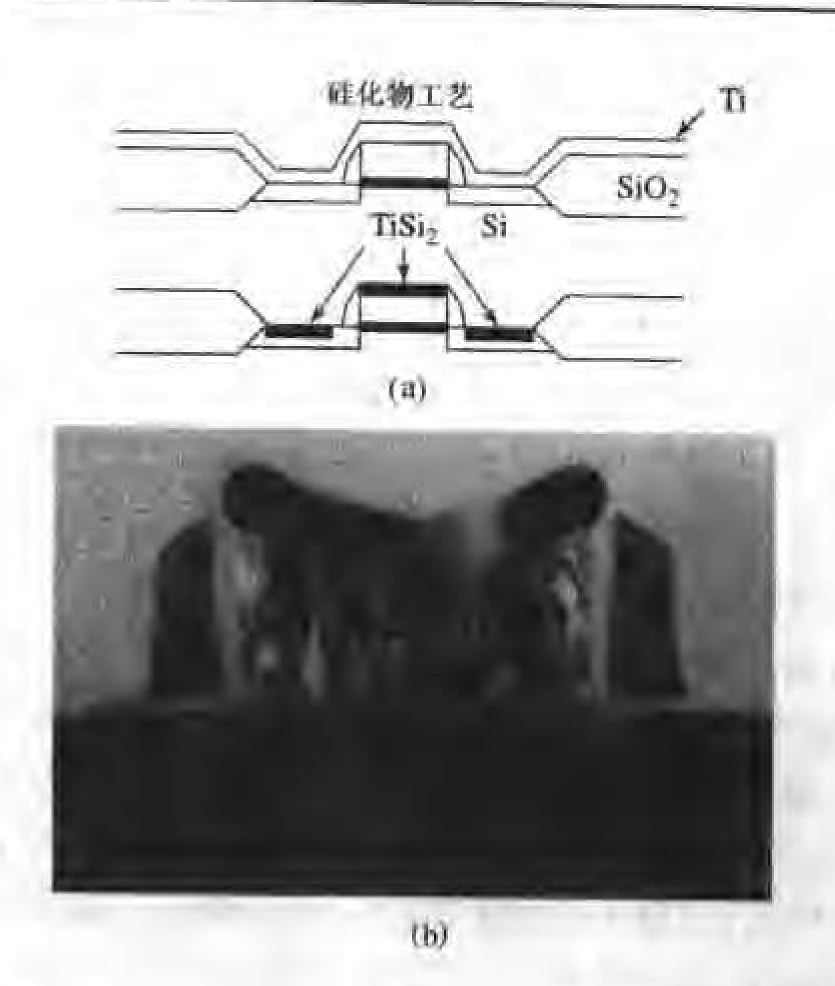


图2.34　(a)在硅化物工艺中，Ti淀积在整个硅片上，退火后在裸露的硅上形成$TiSi_2$。然后采用湿法化学腐蚀将氧化硅上未反应的Ti去除；(b)用硅化物工艺在MOS晶体管栅、源和漏区形成的硅化物的透射电镜的截面照片(引自 *Accurel Systems International Corp.*)

体积。在图片的底部可见覆盖着硅化物的多晶硅连线。非晶二氧化硅层(浅灰色区域)将三层金属铝互连线(颗粒状灰色区域)彼此之间及晶体管和铝互连线之间隔离开。在铝层上面的暗区，有时铝层下面也有，是势垒层，它用做刻蚀终止层，也可阻止互连系统中材料的扩散。这些区域是由重金属组成的，在透射电镜中电子很难穿透，因此在图中看起来很暗。那些连接不同导线间的暗的区域是填充导体间通孔的钨塞。这些钨塞使得电流在不同的金属层间流动，并流向图中底部的器件。多晶硅线之上的暗区是自对准硅化物，用来减小多晶硅互连的电阻。图2.36给出了钨塞连接下面$TiSi_2$/TiN层和上面覆盖的铝层的更详细的透射电镜照片。

随着图形的横向尺寸的减小，垂直方向的尺寸不能按比例缩小得那么快(垂直方向尺寸的缩小是为了减小金属线的电阻和不同互连层上金属线间的电容，因为两者都会影响电路的频率响应)。因此，随着填满金属的连接通孔深宽比的增加，和填充绝缘体的金属线间的间隙的增加，使得工艺制作越来越困难。先进淀积技术，如前面讨论的高方向性高密度等离子体淀积技术已被开发，以辅助填满这些小的空间。

对于更小的图形尺寸和更多的金属层，IC不同区域垂直方向上的高度差已成为重要的问题。随着光刻设备分辨率的提高，它们的聚焦景深不断下降。另一方面，对于复杂的金属化系统，必须在几个不同的深度对电极进行图形化。为了减小这种不兼容性，目前已发展了一些工艺以获得更平的表面。

一种平坦化工艺是在每一步互连工艺后对淀积的介质层抛光，为随后的金属层提供光滑和平坦的表面。这种化学机械抛光(CMP)结合化学和机械的方法去除不需要的材料，能提供

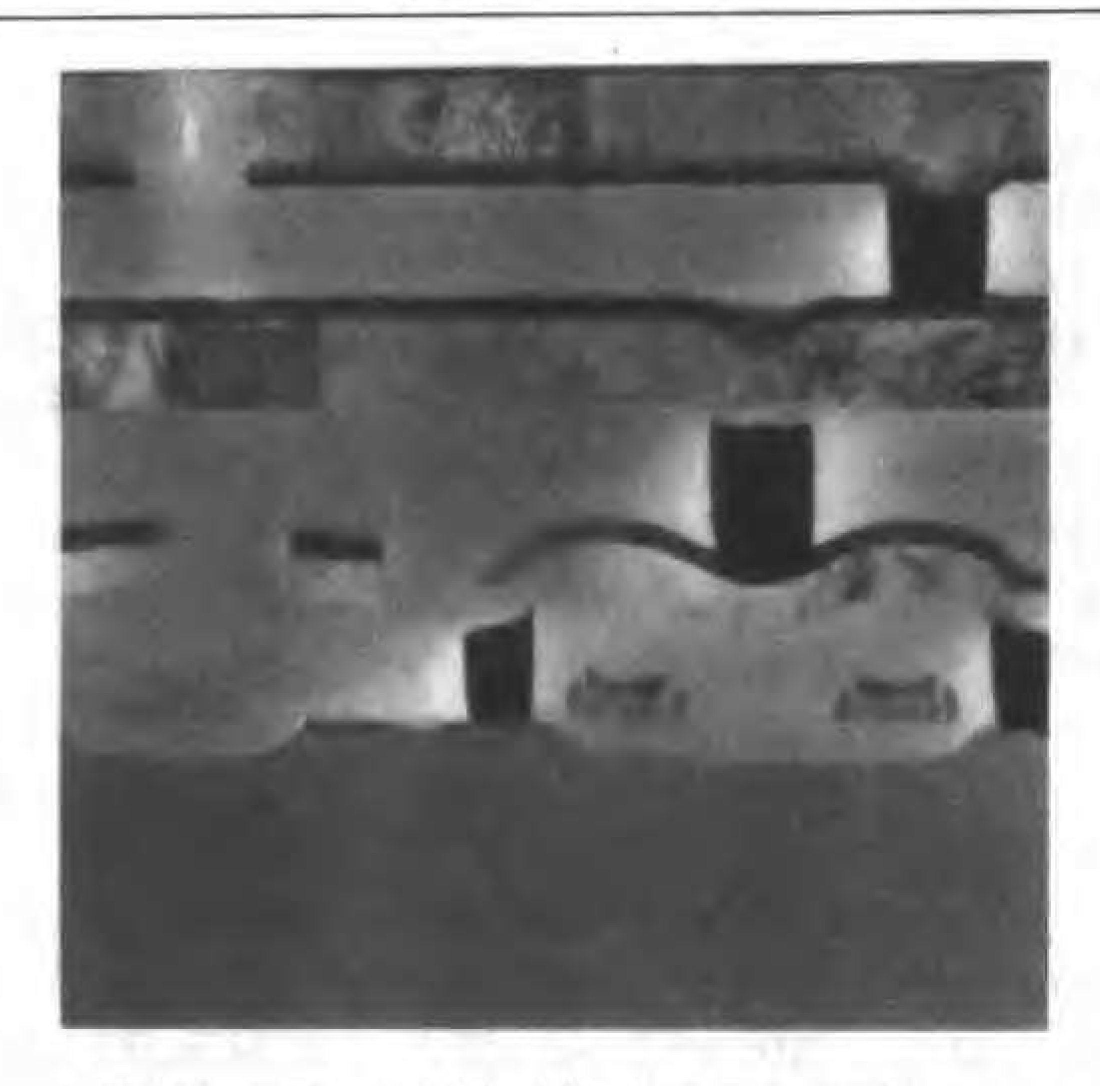


图2.35 三层互连系统的透射电镜截面图。可看到三层金属化铝，相关的势垒层及金属层间填满钨的通道。多晶硅线位于衬底上方(引自 Rudolph Technologies Inc.)

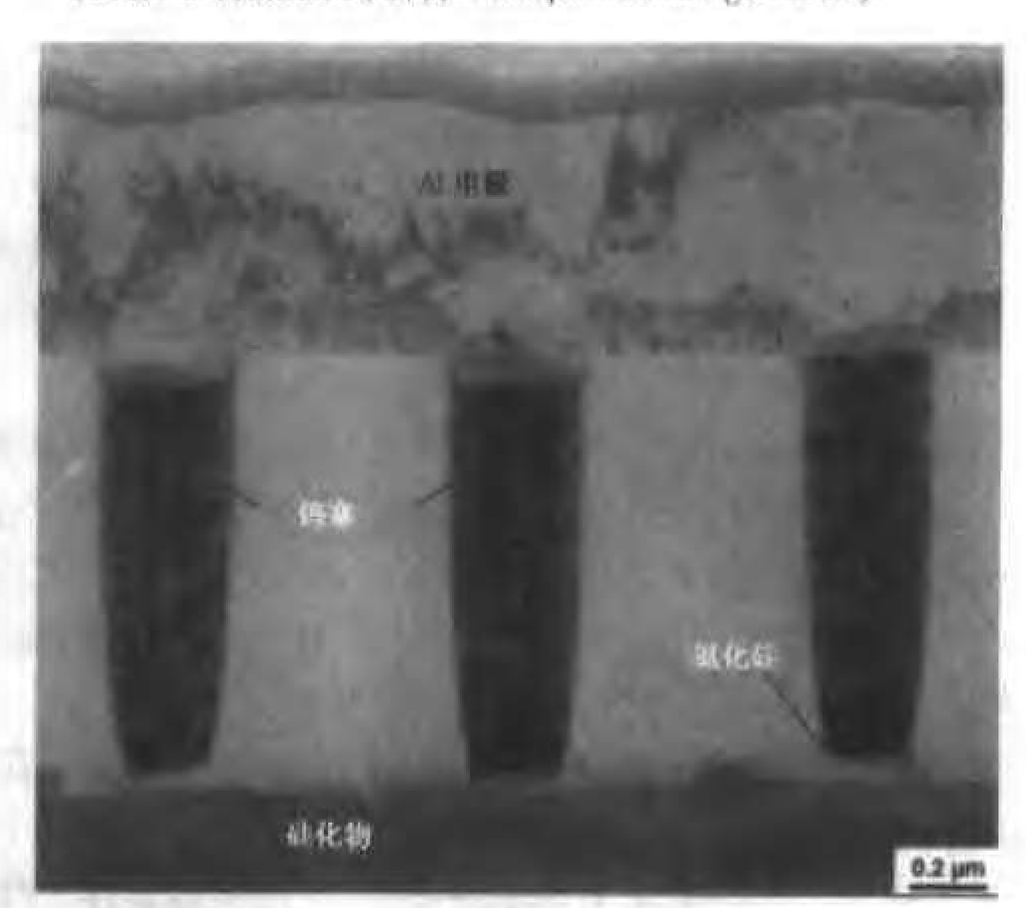


图2.36 透射电镜截面图更详细地给出了钨塞连接下面的硅化物层和上面的第一层金属铝(引自 Accurel Sysytems International Corp.)

非常光滑的表面，而对下面的结构损伤最小。在多层金属系统中，CMP 形成的平坦表面非常有利于形成后面的金属之间的互连。在用高密度等离子体淀积氧化硅填充相邻金属线条之间

窄的间隙时，会形成粗糙的表面形貌，采用CMP进行平面化处理也非常有效(图2.27(d))。CMP工艺能为精细的图形排列提供平坦的表面，但对空间间隔较大的图形之间的表面却不能奏效，因为当同时去除较软和较硬的材料时，在抛光的最后阶段较软的材料会出现凹面。

还有一种不太常用的平坦化工艺，它用CVD方法淀积一层厚的氧化硅，然后采用常规方法在硅片上旋转涂覆一层光刻胶。光刻胶将填满相邻图形间的空间，在局部形成更平的表面。然后用干法刻蚀工艺以几乎相同的速率除去光刻胶和氧化硅。当刻蚀去除所有的光刻胶和某些氧化硅后，氧化硅层具有与光刻胶相似的平整的表面。

铜 随着器件尺寸的减小，对金属化系统的要求越来越严格。例如，我们将在第9章中介绍，系统地减小(按比例缩小)IC中特征尺寸的最常用的方法会导致互连线中的电流密度增加。如果表面的器件尺寸减小K倍，电流也应该以同样的比例减小(表9.1)。然而，如果互连线的截面积下降K^2倍，那么流过互连线的电流密度增加K倍，这将导致在互连线上有更大的电压降，从而外加电压中只有小部分用于驱动IC器件。为了使这一效应降至最低，减小互连导线的电阻率是很重要的。

在集成电路工艺发展的大部分过程中，铝是最主要的互连金属。我们已经讨论过，通常在铝中加少量的硅以阻止它与下面接触的硅发生反应，加少量的铜可以减少下面将讨论的电迁移。但基本材料仍是电阻率为3μΩ·m(体材料)的铝。当互连系统及与其相关的RC时间常数日益成为限制总的集成电路性能的因素时，则需要电阻率更低的金属。铜的电阻率约为1.7μΩ·m(体材料)。尽管将铜与IC制造工艺集成有些困难，但是由于电阻率低，在IC金属化系统中铜还是很诱人的。

一个主要的问题是铜会在硅的带隙中引入深能级，见表1.4。这些深能级增加了漏电流，使得晶体管的增益下降。因此必须十分小心，保证铜没有扩散进入硅，而只局限在金属化系统中。铜可以穿过二氧化硅迅速移动，这使得问题复杂化。因此必须在铜的周围加扩散势垒，这增加了制备工艺的复杂性。另外，对非常细的线条，用各向异性反应离子刻蚀技术很难对铜进行刻蚀；刻蚀反应的产物也不是挥发性的，所以不会离开表面，从而有效地阻止了进一步刻蚀。因此，对铜需要采用另外一种不同于常规的刻蚀金属互连系统的工艺。

在绝缘层(通常是二氧化硅)形成后，先在上面刻蚀出沟槽，然后在沟槽的底部和侧面形成合适的阻挡层。将所有沟槽用铜填满。阻挡层和铜也会淀积在沟槽外的绝缘层上。这些多余的材料用化学机械抛光去除掉，得到平坦的表面以制作下一金属化层。这种所谓"大马士革镶嵌"(Damascene)工艺(名称取自与叙利亚，大马士革有关的镶嵌珠宝制作技术)可用来形成金属线之间的通道(双Damascene工艺)，也可用于制作线条本身。

当总的电路性能受到互连系统RC时间常数的限制时，减小电容和电阻是有利的。因此，那些比常规二氧化硅的相对介电常数低的材料就变得很诱人。在氧化硅中加入氟，相对介电常数可从3.9~4.1降低到3.5。然而，随着氟的增加，氧化硅的稳定性下降，从而限制了添加氟的数量。其他具有低介电常数的材料包括硅、氧和碳的混合物，非晶碳或碳氟混合物等，也可使用介电常数小的有机材料。但是，不论材料发生什么变化，形成互连系统的总的工艺必须保证具有数以百万计线条和通孔的芯片成品率足够高，并能在正常的条件下可靠地工作许多年。

例题 互连延迟

考虑一时间常数 τ_d 为 10ps 的器件，驱动电容为 C，电阻为 R 的金属互连线。互连线的宽度 W 等于 0.3μm，厚度 t_M 等于 0.2μm。金属线与下面的导体用厚为 t_I = 0.3μm 的绝缘层隔开。计算器件和长为 1mm 的互连线的总时间常数，并计算在下列情况下互连延迟与器件延迟相等时的互连线的长度。

(a)金属为铝，绝缘层为二氧化硅

(b)金属为铝，绝缘层相对介电常数为 2.5

(c)金属为铜，绝缘层为二氧化硅

(d)金属为铜，绝缘层相对介电常数为 2.5

已知材料特性如下：

$\rho_{Al} = 3.2\mu\Omega \cdot m$

$\rho_{Cu} = 1.7\mu\Omega \cdot m$

$\epsilon_{r,SiO_2} = 3.9$

利用简单的平行板电容器公式。

解：金属线的 RC 时间常数为

$$\tau = \left(\frac{\rho L}{W t_M}\right)\left(\frac{\epsilon_0 \epsilon_r W L}{t_I}\right) + \tau_d$$

$$\tau = \frac{\rho \epsilon_0 \epsilon_r}{t_M t_I} L^2 + \tau_d$$

(a) τ = 18ps + 10ps = 28ps

(b) τ = 12ps + 10ps = 22ps

(c) τ = 9.8ps + 10ps = 19.8ps

(d) τ = 6.3ps + 10ps = 16.3ps

互连延迟等于器件延迟时的互连线长度

(a)0.74mm

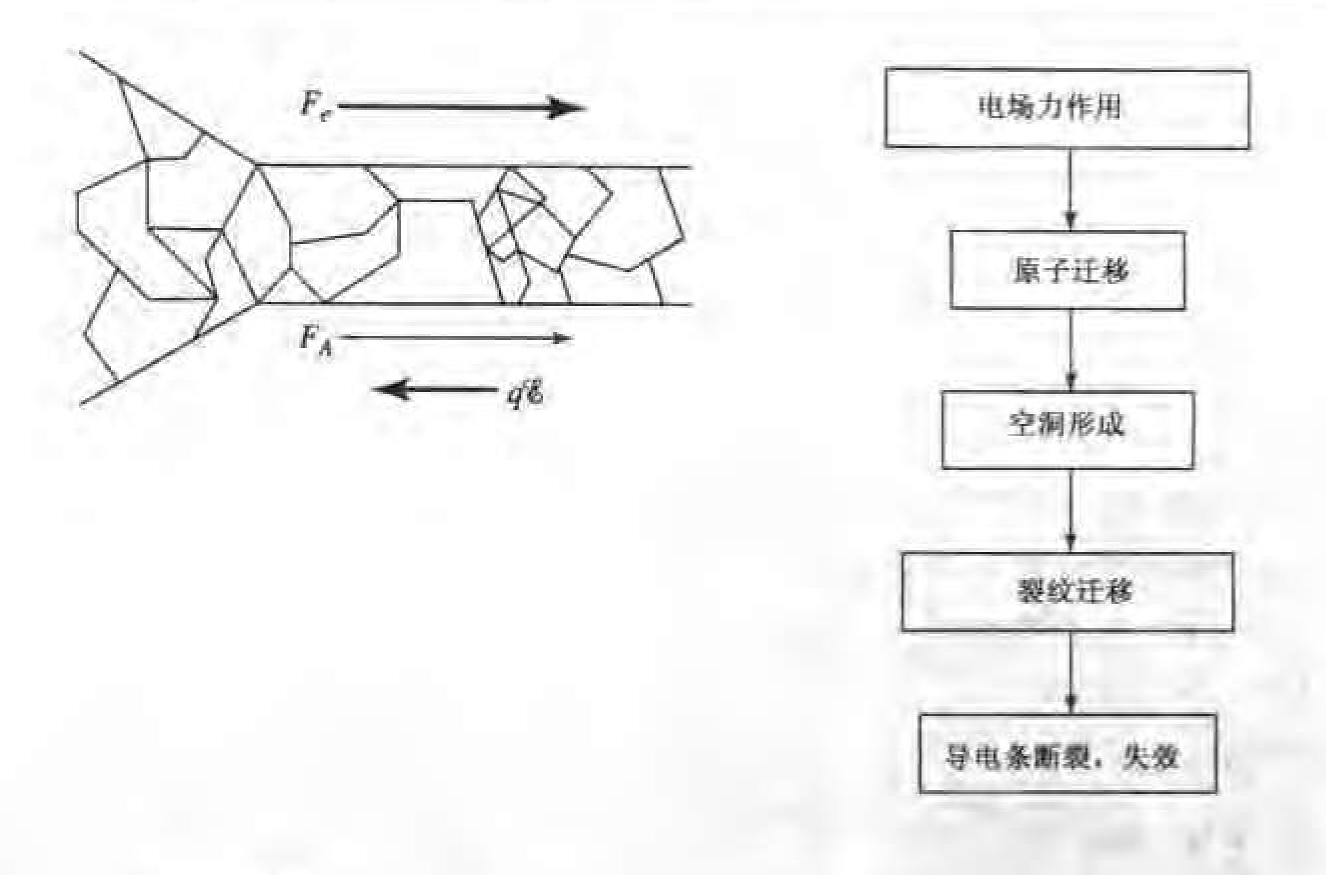
(b)1.01mm

(c)0.92mm

(d)1.26mm

电迁移 电迁移会影响 IC 互连的可靠性，即在正常工作了成百上千小时后，互连线会产生断裂而造成电路失效。电迁移是导电材料的原子发生运动，可动载流子与原子晶格间的动量交换的结果。金属中运动的电子与原子发生碰撞，将原子推向正电极(图 2.37)。结果金属在这一电极附近堆积，而在导体的其他部分尤其是在多晶金属膜的晶粒间界区域耗尽。材料的迁移最终引起薄膜中的空洞和互连线断裂。电迁移在高电流密度和大温度梯度处进行得更快。

电迁移是铝互连线出现的主要问题，要求设计的电流密度值通常要小于 $10^5 A\ cm^{-2}$。在铝中增加少量的其他金属如铜以阻止铝原子沿晶粒间界的运动，能减小电迁移。增加 2% ~ 3% 的铜就能将长时间的电流处理能力提高两个数量级而不明显增加薄膜的电阻率。与铝相比，



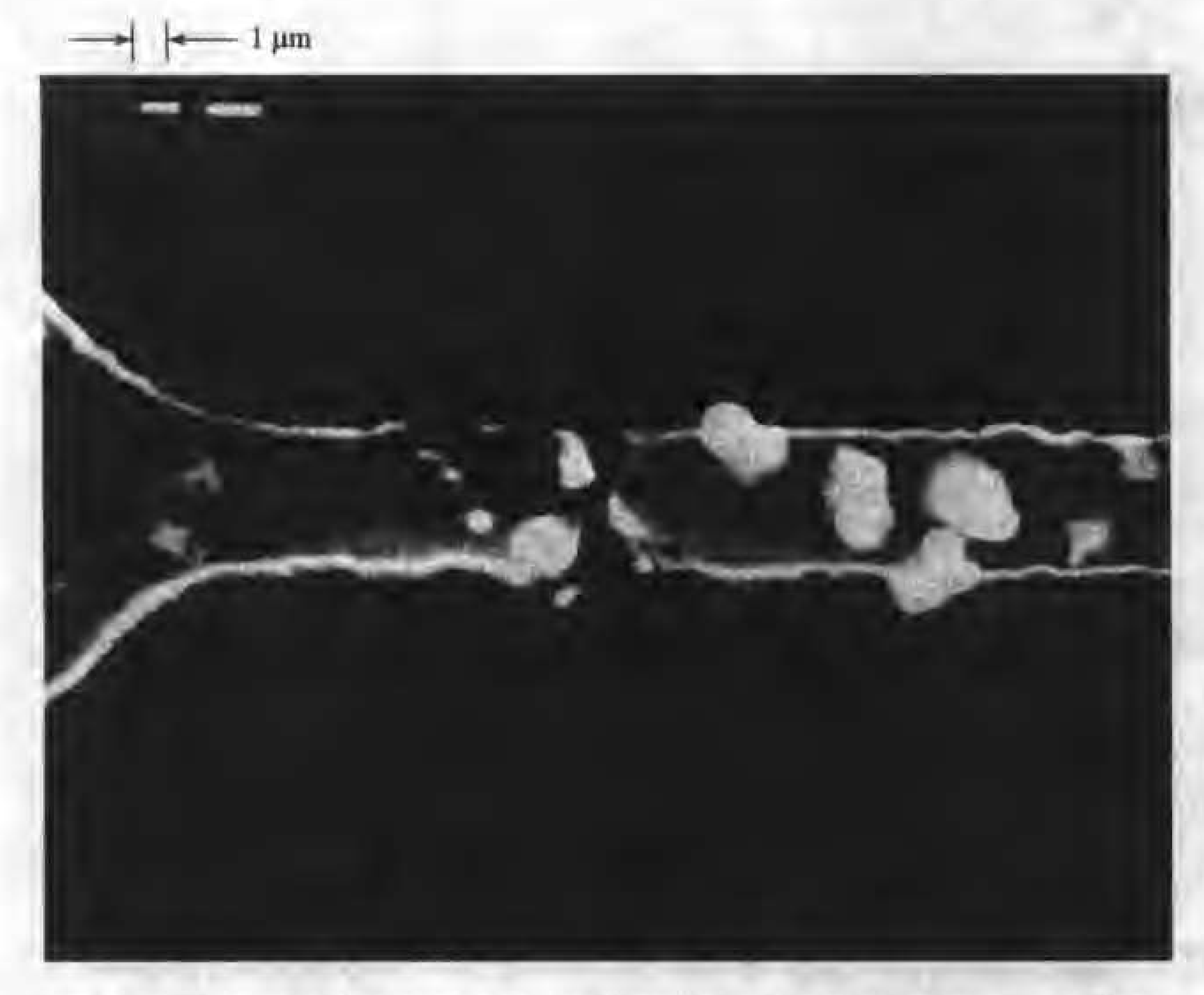


图2.37 导电条中的电迁移机制。左上图给出了电子流 F_e、静电力 $q\mathscr{E}$ 及形成的原子流 F_A 的方向。扫描电镜像表明断裂处左侧空洞的形成，和断裂处右侧金属积累形成的小丘(下图)。电迁移造成的失效过程在右上图给出[38]

铜的电迁移较难发生，高的电流处理能力是用铜取代铝的另一个原因。另外，可用难熔金属如钨做金属线，因为它也很难发生电迁移。

在淀积金属互连层并图形化后，将硅片放置在低温炉中(对铝约为450℃)合金可提高互连系统中金属的界面特性，并保证良好的欧姆接触。这一热处理过程也提高了 Si-SiO_2 界面的质量。当互连线图形化后，IC 芯片的加工就完成了。

除了电迁移和电阻率的问题，选择互连材料还要考虑以下几点：(1)与n型和p型硅都能形成欧姆接触的能力，(2)电路制作完毕后与硅接触的稳定性，(3)与硅和二氧化硅的附着力，(4)采用常规的光刻和刻蚀（主要是干法刻蚀）或抛光技术进行图形化的能力，(5)抗环境腐蚀的能力，(6)与合适的封装材料进行连接键合的能力，(7)能覆盖IC工艺中的台阶，(8)在不降低已有器件性能的条件下被淀积的能力。没有一种金属能够满足以上全部要求，铝及其合金能够满足以上的大部分要求，因而被广泛使用了很多年。然而，随着对互连的要求变得日益苛刻，铝的局限性（尤其是电迁移）变得更显著，导致了对替代金属的研究和铜的广泛使用

2.7.2 测试和封装

硅片制备工艺完成后，需要对IC进行电学测试以决定哪部分能正常工作，只需对正常工作的芯片封装。随着集成电路复杂性的增加，在计算机控制下进行的检测已变得非常困难。在电路设计和版图设计过程中，如何便于电学测试已成为重要的考虑因素。加入适当数量的外加电路有利于进行有效的检测。在设计芯片时，要保证关键的内部电压可从外部测量，以确定电路是否能正常工作。

在进行了初步的功能测试后，用具有锋利的金刚石针尖的设备（金刚石划片机）沿着硅的解理面将硅片划分为单个电路，即芯片。其他划片工艺包括将硅片锯开或用激光将硅片融化至一定深度后分开。在最简单的封装方法中，将每一功能芯片的背面焊在管壳上，半导体芯片表面的压焊块与管壳的引脚用导线连接或键合（图2.38）。最后，管壳用保护性的陶瓷、金属或塑料壳密封。还需要进一步对电路进行电学测试以确认其仍能正常工作，通常还要测量频率响应。

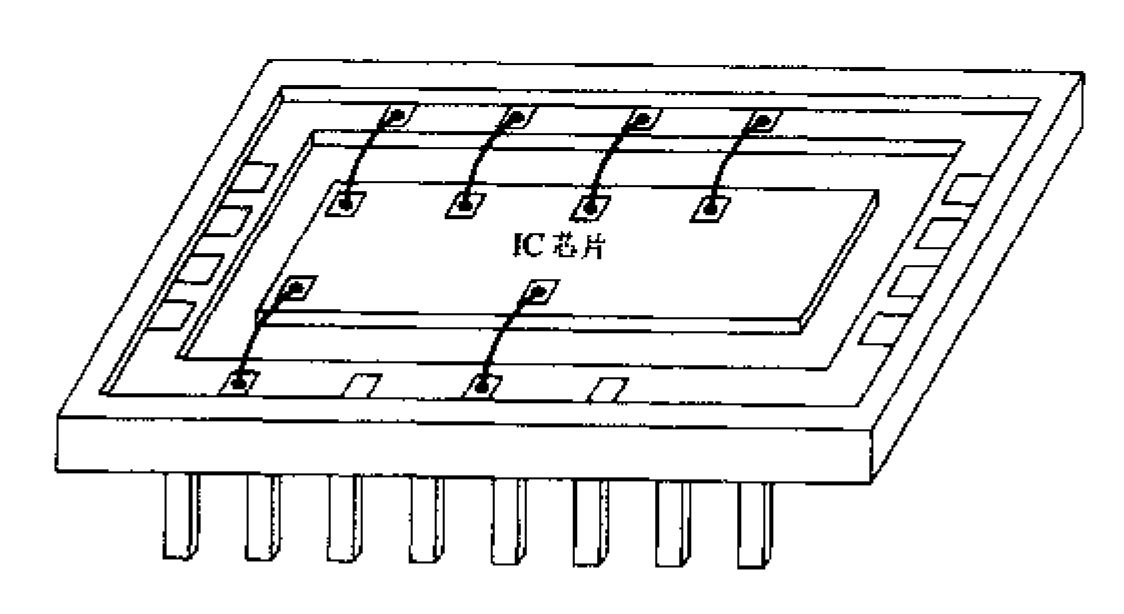


图2.38 IC芯片固定在管壳上，用导线与管脚相连

当设计包含更复杂IC芯片的系统时，芯片间用于通信的互连线数量增加，需采用其他封装技术。为得到更多的互连线，更高的可靠性和封装密度，出现了含有几层金属互连线的陶瓷衬底。IC芯片面向下与这些衬底键合，使IC芯片上的压焊块直接连接到陶瓷衬底上相应的金属块上。通过融化预先在IC压焊块上形成的焊料凸点，实现所有引线同时键合，这种工艺称做倒装焊。倒装焊能够同时连接数百个引线，并能连接靠近芯片的中心区域。

功耗是设计高性能IC和选择合适的封装技术需要考虑的一个日益重要的因素，常规的封装技术限制了芯片的功耗不能超过几个瓦特；更精致的封装和冷却技术（例如采用热沉甚至风扇）可允许芯片功耗达到几十瓦特。随着器件尺寸的减小，在给定尺寸芯片上产生的热量

最好能保持不变,但通常都会增加。大的芯片上功耗增加得更多,因此功耗(及所要求提供的功率)会限制 IC 芯片的复杂性(尤其在便携应用方面),并最终导致优先选择某种类型的电路技术。

双极电路一般比 MOS 电路消耗更多的功率,单沟道 MOS 电路比互补型 MOS(CMOS)电路消耗更多的功率,CMOS 电路的直流功耗很小。然而,即使在 CMOS 电路中,功耗也是一个越来越大的约束条件。尽管大部分 CMOS 集成电路中的直流功耗很小,对互连线充电和驱动晶体管的交流功耗也可以很大。交流功耗为

$$P_{ac} = \frac{1}{2}CV^2f \tag{2.7.1}$$

式中,C 为电容,V 为电压的变化(几乎与 CMOS 电路的外加电压相等),f 为频率。减小外加电压可降低功耗,但是增加频率(高性能电路所要求)又大大地增加了功耗。各种不同的 IC 技术在以后的章节中要进一步讨论

2.7.3 污染

随着重复单元尺寸的增加和特征尺寸的减小,由漂游的颗粒或“灰尘”引起的损伤变得日益严重。芯片尺寸越大,能容忍的颗粒数目越少。特征尺寸越小,能破坏芯片的灰尘颗粒的尺寸也在减小。集成电路的制造设施要保证靠近硅片的颗粒数最小。尽管造价很高,10 级超净间(每立方英尺 10 个颗粒)和 1 级超净间(每立方英尺 1 个颗粒)已被广泛使用。

2.8 化合物半导体工艺

本章我们重点讨论了发展很成熟的硅工艺,绝大多数的半导体器件都是用硅工艺制作的。但是,在某些特殊应用中,必须使用其他半导体材料。某些化合物半导体具有更好的输运特性,可用于制作非常高速的晶体管。由元素周期表中Ⅲ族和Ⅴ族元素组成的材料,如砷化镓,可用在要求有高载流子迁移率的专用集成电路中;Ⅱ-Ⅵ族材料用得较少。另外,直接带隙化合物半导体如 GaAs 和 GaN,被用在发光二极管(LED)和半导体激光器等光发射器件中,以实现高效率光发射。硅具有间接带隙,不适合于这些方面的应用。

由于在高速器件和光电子器件中的重要性,下面简单介绍一下与化合物半导体有关的一些加工工艺。重点介绍 GaAs,因为它是最广泛使用的化合物半导体材料,并以此为例说明硅和化合物半导体在加工工艺上的一些差异。

尽管 GaAs 特别适用于某些专用领域,但有限的需求量限制了 GaAs 工艺的发展,对获取高集成度 GaAs 集成电路的关注也比硅 IC 要差得多。因此,GaAs IC 中晶体管的数目比硅 IC 低几个数量级。化合物半导体工艺的重点在于如何利用非常小的特征尺寸获得高的性能,或如何实现硅不能实现的功能,而不是追求高的集成度。

化合物半导体工艺沿着与硅工艺不同的路线发展,形成了另外一套加工方法。然而,对硅加工工艺的投资如此之大,远远大于只有小批量生产的化合物半导体加工工艺,许多化合物半导体的工艺是向硅工艺靠近的,这也带动了硅工艺的不断发展。反过来,当硅工艺的热开销受到更多的限制时,不再倾向使用厚的热生长氧化硅,并且硅工艺也正在吸收化合物半导体加工的一些特点。因此,以前在硅和化合物半导体加工中存在的一些差异逐渐变得不太重要了。

因为GaAs是最常用的化合物半导体，下面的讨论中将重点介绍。GaAs有几个基本特征使得对它的加工与硅不同。首先，GaAs是化合物，Ga和As含量稍微不同就会偏离理想的化学计量比，产生的缺陷将降低其电学性能。其次，在GaAs上很难形成稳定的氧化物，使其表面性质很难控制。

晶体生长 与硅一样，GaAs也可用直拉法生长。然而，生长化合物半导体更复杂，因为半导体的化学计量比及晶体质量必须很好地控制。Ga和As的含量稍微不同就会产生具有电活性的缺陷。与Ga相比，As很容易挥发，所以As在融化状态下蒸发很快，从而形成富Ga的晶体。为了阻止As从熔融材料中蒸发，晶体生长设备可放置在包含高砷蒸气压的密封腔体内。这种方法复杂且危险，因此一般在融化GaAs上面覆盖另外一层液体以阻止砷的蒸发。这种液体密封直拉法(LEC)在大直径GaAs晶体生长中广泛采用。

GaAs单晶也可用类似于硅的区熔法进行生长。水平Bridgman工艺有时用于小直径硅片的生长。在晶体生长过程中晶锭是水平的，因此不是圆柱形的，从晶锭上切割下来的圆片更像D形而不是圆形。与硅一样，随着直拉工艺控制技术的提高和圆片直径的增加，区熔法已不太流行了。

因为控制多元素材料的生长要比控制单元素材料生长困难得多，GaAs和其他化合物半导体的生长技术落后于硅。通常晶圆片的直径明显要小，且成本很高。化合物半导体也比硅容易碎，机械加工时必须十分小心(自动化加工时尤其重要)。与硅相比，要更仔细地控制加热和冷却速率以使热冲击降至最低。化合物半导体的化学稳定性不如硅，许多对硅不起作用的化学物质会对其造成侵蚀。

高电阻率 GaAs比硅的带隙宽，所以它的本征电阻率明显要高。根据等式(1.1.25)，为了得到最高的电阻率，Fermi能级应位于禁带的中央，以减小自由载流子浓度。尽管将硅的Fermi能级控制在带隙中部有困难，但GaAs可在生长时掺入高浓度的特种杂质，如Cr，它在靠近带隙中部产生深能级。Fermi能级于是被钉扎在带隙中部，远离两个带边形成半绝缘材料。

最终得到的高电阻率减小了衬底与上面器件间的垂直电容。与硅器件相比，这是高性能GaAs器件的一个重要优势。对硅器件，有源区与衬底只通过有限厚度的pn结隔离，使得器件与衬底间的电容不能被减得很小

尽管硅上的绝缘氧化物可以将金属互连线与导电衬底隔离，但氧化层的厚度有限，使得衬底与互连线间的电容很大。对于GaAs，厚的高电阻率衬底也会减小这一电容分量。

对硅器件，生长的绝缘氧化物可用于相邻器件间的横向隔离。GaAs衬底几乎是绝缘的，所以通过离子注入将杂质注入器件有源区，器件间则保持为半绝缘的区域即可实现器件间的隔离。如果半导体层生长在整个半绝缘衬底上，将器件间的半导体层刻蚀掉就可以得到很好的隔离。如果用氢离子注入某些特定的区域，破坏它的晶格且不进行退火消除损伤，GaAs的电阻率还会进一步提高。这样得到的高电阻率可用于器件间的横向隔离。尽管这些技术为相邻器件间提供了直流隔离，因为GaAs的介电常数高，器件间的横向电容仍会很大。

外延 因为其他化合物半导体可在GaAs上外延生长，相对硅而言，对化合物半导体精确可控的外延生长技术的研究进行得更为广泛。因为化合物半导体种类很多，将不同带隙的材料结合在一起(提高器件的电学性能)，而晶格常数仍然保持不变(限制应力和晶格缺陷)是可能的，如图2.39所示。例如，$Al_{1-x}Ga_xAs$的晶格常数几乎与GaAs相同而带隙要大得多。这

样,可用与讨论硅时类似的外延技术在 GaAs 上生长无应力 $Al_{1-x}Ga_xAs$ 外延层(原子比为 $1-x$[2] 的 Ga 被 Al 代替,所以组分由 $Al_{1-x}Ga_xAs$ 表示。纯 AlAs 是间接带隙材料,所以 Al 的比例通常受到限制,以保证其为直接带隙)。器件的不同部分采用不同的材料,就能利用两种材料带隙的差异构造有用的电子和光电子器件。利用界面处能带不连续性构造的这些异质结器件,将在第 4 章和第 7 章中简单介绍。对异质结器件进行概述后,我们将集中讨论硅和 $Si_{1-x}Ge_x$ 合金组成的异质结器件,这不仅因为其重要性,也考虑到它们与硅集成电路工艺相近(注意 $Si/Si_{1-x}Ge_x$ 系统的晶格是不匹配的,导致的应力将限制外延层的厚度)。

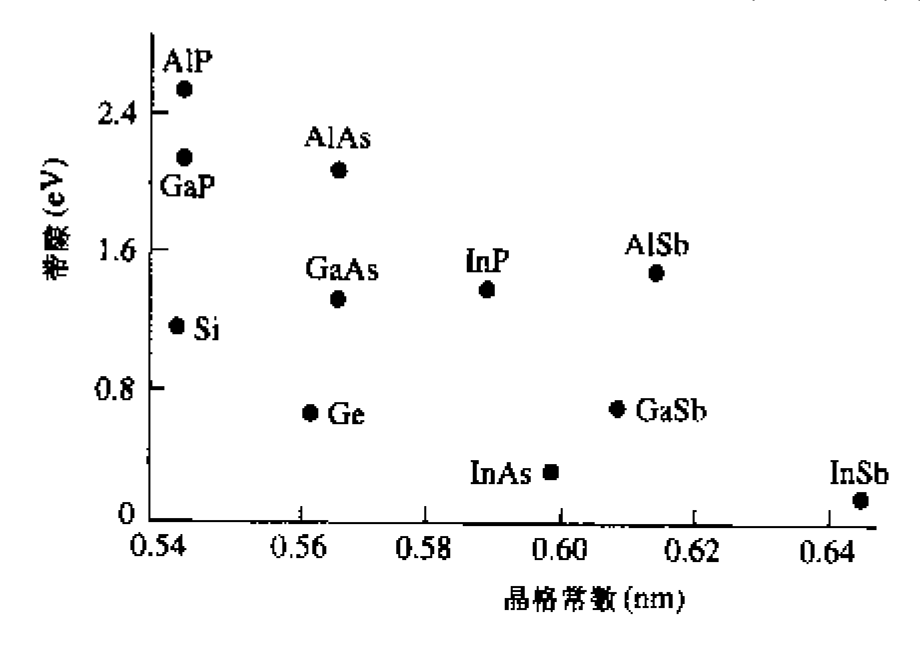


图 2.39 将两个不同带隙的半导体相结合能提高器件的电学性能,但是,晶格失配和相应的应力限制了可用的异质外延材料的组合

因为异质结器件外延层间的界面强烈地影响其电学特性,控制不同材料间的过渡区是十分必要的。已经研究出在几个纳米范围内控制界面的方法,这比控制常规硅器件的界面要精细得多。反过来,化合物半导体在外延后的加工更为困难。

在制作由不同半导体材料层组成的器件结构时,某些材料层必须经常从器件的部分区域中去除而不能对其下面薄的其他材料层有影响。为了获得高的选择性,对 GaAs 及其相关材料的湿法化学腐蚀工艺的发展已经很成熟。

电介质 GaAs 和其他化合物材料面临的一个重要的问题是缺少稳定的热生长的氧化物来进行半导体和上面金属层之间的隔离,并且也不能提供电学稳定的半导体界面。在第 8 章中将会看到,在硅上生长的 SiO_2 层将饱和大部分硅表面的断裂键。这一氧化层减小了硅表面处禁带中允许态的数目,而硅表面是由周期性晶格结构中断形成的。GaAs 中没有类似的氧化物,因而这些表面态的存在使得对界面特性的控制非常困难,可能对器件的性能起主要作用。可在 GaAs 上生长较宽带隙的薄外延层材料如 $Al_{1-x}Ga_xAs$,以提供更稳定的界面。

GaAs 和其他化合物材料缺乏稳定的热生长氧化物也使得在表面选择小区域进行掺杂的过程变得复杂化,CVD 方法制备的电介质在化合物半导体中被广泛使用。等离子增强化学气相淀积的方法尤为普遍,因为等离子淀积温度较低,可使化合物半导体组分的蒸发降至最低。淀积的典型的电介质是氧化硅和氮化硅,通常利用材料组分可调的 PECVD 方法进行制备。

---

2 译者注:原文误为 x。

掺杂 在一定范围内，杂质可以通过气相扩散掺入到 GaAs 中。与硅类似，目前已经越来越多地采用离子注入对器件的选择区域进行掺杂。与硅工艺类似，也可用光刻胶来限定杂质注入的区域。

然而，许多杂质在 GaAs 中的固溶度低，难以获得高浓度的电激活杂质粒子。这就限制了 GaAs 器件只能用于那些其性能远远优于硅器件的场合，这就大大限制了 GaAs 器件的种类。另外，形成好的欧姆接触需要高的掺杂浓度，这也是有限的固溶度难以实现的。

因为 GaAs 的热稳定性较差，通过退火去除离子注入形成的晶格损伤和激活杂质原子比硅困难。而且，当温度高于600℃时，As 趋于从 GaAs 中蒸发。如果在这一温度范围内需要对离子注入进行退火，必须淀积一层材料对表面进行很好的保护，以防止 As 蒸发和表面化学剂量比发生变化。低温下(~300℃)等离子增强 CVD 制备的非化学计量比的氮化硅层可阻止 As 的蒸发，而氧化物则不太有效。另外，如果不对表面加以保护，可在 As 的过压蒸汽中进行退火，但满足该工艺要求非常困难。

互连 一旦杂质原子就位并被激活，就可淀积另外的介质层将上面的导体和下面的有源器件区隔离。介质层形成后，与硅器件一样，要开接触窗口并淀积金属层。为了与 GaAs 形成良好的电学接触，通常需要采用多层金属化系统。过去，与硅工艺中使用的较简单的金属化系统相比，采用多层金属化系统是一个缺点，但是现在硅的金属化也变得更复杂了，而且在硅和主要的金属化材料铝或铜间还要加一中间层。

与硅器件一样，为了获得高性能，必须十分注意互连系统，要获得最好的性能，在互连系统中就必须采用低电阻率的金属和低介电常数的介质。因为 GaAs 集成电路的集成度较低，金属化系统更灵活一些。除了铝，金也经常用做主要的导体。因为金不能形成欧姆接触，必须在金与半导体间加入一层中间金属层。

为了减小电容，金属和器件间需采用低介电常数的介质；与硅集成电路类似，有时采用有机材料。在极端的情况下，金属图形化后，金属和器件间的介质可用湿法化学腐蚀去除。尽管金属与半导体器件间的空气的相对介电常数为1，但这样的"空气桥"很脆弱，只适用于专用的和小规模集成电路。

要充分利用化合物半导体器件的高性能，封装就不应明显降低它们的指标。对高频器件必须采用特殊的具有极低电容和电感值的管壳。对光电子应用，封装有不同的要求。如果要让芯片上的光透出来，部分管壳必须是透明的，同时还要能对里面的电子器件起到足够的保护。光线也必须经过反射和聚焦以便离开芯片和管壳，而不是被芯片的不同区域或封装材料重新吸收。

2.9 数值模拟

2.9.1 模拟的基本概念

随着集成电路工艺和器件变得越来越复杂，求解与器件物理或工艺有关的解析方程变得更困难了。当器件尺寸更小、结更浅时，必须考虑二级的物理和电学效应。同时，对器件和工艺性能的预测也变得更复杂了，进行有价值的实验验证也越来越困难。为了清楚地揭示被研究的物理效应，对实验进行控制的难度甚至可能会超过设计实验的难度。随着加工设备越来

越昂贵,实验的成本也在不断增加 同时,自动化程度的提高使实验的灵活性下降,实验也变得很难做 由于可能出现的变化实在太多,实验上已无法研究所有相关变量组合在一起的结果

相反,随着计算机的功能日益强大,数值模拟越来越吸引人,人们正利用模拟来逐渐替代各种各样的实验 随着模拟精度的不断提高,物理制造过程和器件的电学特性都可以用数值方法模拟 由于数值方法具有快速模拟许多工艺变量和器件变量的能力,因此能够确定出优化的条件参数,而实验工作只需集中在数量有限的变量上 另外,与以往比,现代器件要求更精确地了解器件的结构参数和掺杂分布 器件的性能越来越依赖于这些参数,也越来越依赖于更精确的理解这些结构中的载流子在电场作用下的行为

模拟技术可以是严格的数值方法,也可以将数值方法与解析表达式相结合。如果存在解析解,则能帮助设计人员更好地了解工艺过程和器件物理,同时减小所需的计算时间。然而,如果采用更精确的描述工艺或器件的表达式(例如,利用更复杂的扩散方程式(2.5.15)代替简单的 Fick 定律表达式(2.5.7)),通常只能用数值方法求解这些方程。

现有数值模拟工具的功能在不断增强。工艺和器件模拟器中结合了更实际的物理模型,可以模拟二维或三维的情况。随着器件尺寸的减小,器件边界的影响范围已扩展到占据器件相当大的比例,会明显改变器件的制作和器件的特性。相邻器件单元间的横向效应也变得越来越重要。因此,现代的工艺和器件模拟中能包含二维效应是很关键的。但是,随着模拟维度的增加,数值计算能力也需要迅速增加,一般来说多维模拟器不能像一维模拟器一样包含那么多的详细的物理模型。鉴于在拐角处的效应也变得越来越重要,这就要求进行三维模拟计算,其复杂性也在不断提高。当包括二维和三维效应时,所要求的计算能力迅速提高,如果可能的话,用于数值估算的解析模型是非常诱人的。当考虑二维效应时,应该考虑靠近边界处的图形的形状和杂质的分布。例如,图形较深时,化学气相淀积层的形状非常重要,在建模中应予以考虑。刻蚀精细的图形也要求考虑边界的形貌,即二维效应。

计算机模拟程序根据物理模型和数值参量进行计算,它的精度只由使用的参数的精度决定。进行精确的模拟要求对模型中大量的参数有精确的了解 了解这些参数本身就是一个重要的挑战。放置在模拟程序中的缺省值是模型编制者从文献或实验中选取的最好的猜想值。当可以得到更精确的实验值时,这些参数可以被修正,许多程序都允许使用者来修正数值参量。另外,除了对参数的有限了解外,还需对特定的加工设备中的一些参数进行修正;例如,氧化与氧的压力有关,在高纬度下氧的压力明显要低一些(例如,科罗拉多的气压比加利福尼亚的低)。

2.9.2 网格

数值模拟是基于对器件感兴趣区域中的许多点或节点来计算其特性的。这些特性是根据时间的变化和相邻节点的值计算出来的。节点是叠加在器件上的网格的交叉点 网格的产生是计算机模拟中最困难的任务之一。在器件的有源区,参数随位置迅速变化,网格必须足够细才能通过计算得到详细的特性;而参数随空间位置变化不大的区域,不需要过多的计算时间。网格的间距取决于物理参量的大小,也与运动粒子的特征长度有关 同时模拟运动距离不到一个微米的粒子(例如杂质原子)和运动距离达几百微米的粒子(例如点缺陷)是非常困难的,因为两者的空间尺度相差太大。在模拟过程中当区域间的边界移动时,例如氧化或硅化物的

形成过程,确定网格尤其麻烦。区分不同材料的边界不再与节点保持相同的距离,节点可能会移动到不同的材料中去。在同一个模拟中,如果必须同时考虑时间常数变化很大的各种粒子的情况,确定时间增量也很困难。

在一维模拟中,选择节点的间距与预期的空间变化的尺寸相一致。网格线间距可调,这样在器件结构或电学参数变化较大的区域节点密一些;而在特性缓变的区域节点间距大一些以节约计算时间。在二维模拟中,划分网格更复杂。为了避免过长的计算时间,节点数目应该越少越好。没有必要采用一套规则的等间距的网格线,而一般采用非均匀网格使得节点集中在物理特征或电学特性变化较快的区域。随着器件尺寸的减小,器件不同区域的实际形状对器件的特性有更大的影响,因此能对任意几何形状进行模拟变得尤其重要。然而,处理任意形状时产生网格变得更困难。

在使用典型的二维模拟器时,使用者从确定非均匀的矩形网格入手。矩形网格的线条间距不均匀,在特性变化剧烈的区域放置更多的节点,一部分网格线可以去掉。然后,通过加对角线将修正过的粗矩形网格自动转换为三角形网格。对于三角形网格,一个节点是三条或更多的线段的交点。每一个三角形的面积被构成该三角形的三个节点共享。模拟器可在参数变化剧烈的区域增加网格的密度。例如,在 MOS 晶体管的源漏区边界处掺杂变化快的地方,网格可以进一步细化。在器件模拟中,对某一特定的偏压求解 Poisson 方程后,可以在电位或电荷密度变化剧烈的区域进一步细化网格。

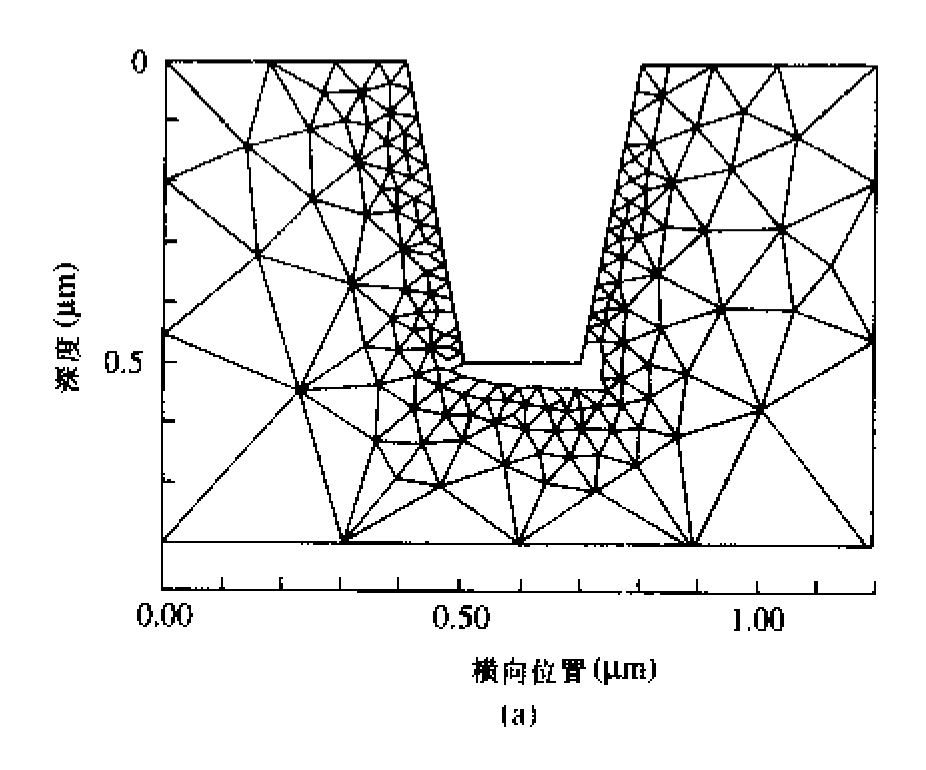
在加工过程中,随着结构的变化,网格也必须变化。例如,在氧化过程中,Si-SiO_2界面会发生移动通过现有的节点。这些节点必须从硅中除去,而另外的节点必须加入到不断生长的氧化层中。同样,淀积和刻蚀工艺也要求节点不断地增加和去除。这些修正由模拟程序本身自动进行。

网格必须足够细才能反映出结构或杂质分布(工艺模拟)或电位和载流子浓度(器件模拟)的剧烈变化。网格应能适应器件形状的变化,分辨出被模拟结构的几何特性。然而,随着节点数目的增加,计算时间增加很快。如果用 N_p 代表节点数,计算时间近似按 N_p^{α} 增加,α 介于1.5至2之间。网格排列的不规则性也会引起收敛问题。图2.40[19]给出了用于非矩形结构的三角形网格。如果在这些非矩形结构中采用矩形网格将会使节点数大大增加。

2.9.3 工艺模型

工艺模拟器总的来说是用与硅片上实际进行的工艺操作相同的顺序,计算每一个工艺步骤造成的变化。除了预测最后的杂质分布以外,工艺模拟器还能计算某些简单的电学参数。计算得到的杂质分布也可输入器件模拟程序以进行更广泛的器件模拟。

工艺和器件设计同时需要更精确的模型,模拟器也在从研发领域逐渐过渡到面向制造业的工程领域。在这种环境下,由于需要模拟出器件制造的全部工艺过程,并且最终给出器件的特性,所以模拟的复杂性不断增加。为了能模拟很长的工艺流程,首先需要研发出包含数量有限的相关工艺步骤的短模块。然后,将这些模块连在一起就可以对整个器件制造过程进行模拟。在模拟时,首先确定描述一系列工艺步骤的模块,然后将这些模块连接在一起(通常在图形界面上移动代表单个模块的图标即可),模拟完整的工艺流程。这样,可以在每一模块中改变参数而无需重写全部规格说明。



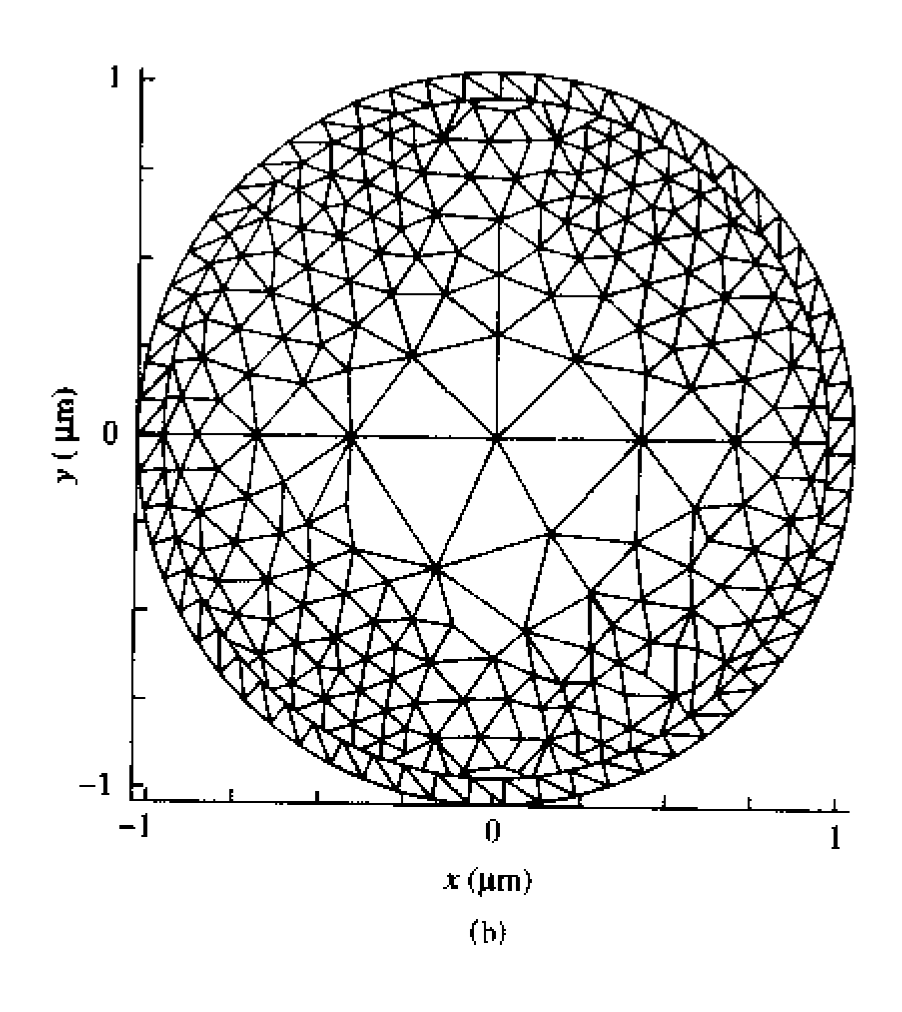


图 2.40 在非矩形结构中划分的三角形网格

(a)具有倾斜侧壁的沟槽结构;(b)圆形结构[19]

当模拟工作进入到生产环境,还必须考虑预期的工艺变化对器件特性的影响。为了得到经济上可以接受的成品率,在设计工艺和器件时,必须考虑单个工艺步骤的统计分布对器件性能的影响。即必须选择工艺和器件参数,使它们对任意工艺参数的正常浮动不过度敏感。模拟为研究工艺变化对器件特性的影响提供了一种有效的方法。

SUPREM 一种广泛使用的工艺模拟程序称为 SUPREM(斯坦福大学工艺工程模型的简称)[20,3]。SUPREM 程序的输入是对单个工艺步骤的描述。输入语句说明了有关扩散、氧化、

3 有几个商用的 SUPREM 的派生和变化的版本

注入、淀积和刻蚀工艺的一系列时间、温度、气氛和其他参数。输出的结果是硅衬底和二氧化硅或多晶硅等覆盖层中的杂质分布。SUPREM 程序的基本结构如图 2.41 所示，工艺步骤可以单独模拟或连续模拟，模拟一个工艺得到的杂质分布又可作为下一个工艺的输入。这个程序包含有关非线性扩散、氧化过程中的杂质分凝、固-气相界面的蒸发、在生长的 SiO_2 层下面 Si-SiO_2 界面的移动、扩散过程中杂质的聚集、浓度增强氧化、外延和离子注入等详细的模型，以及几个可以用解析方式求解的超出了一级效应的模型。复杂的点缺陷模型也包含在内，因为点缺陷对浅结的杂质扩散很重要。

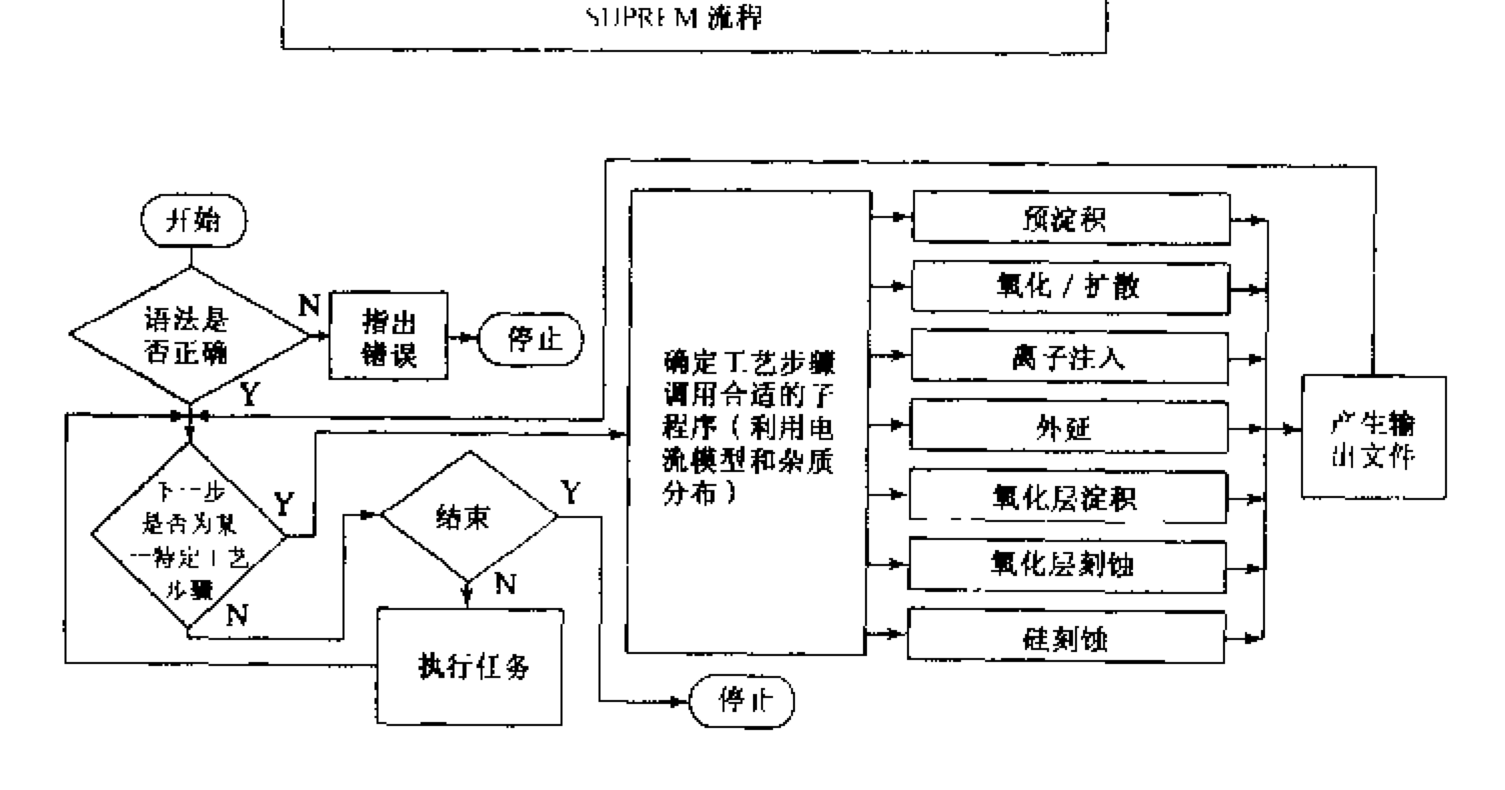


图 2.41 SUPREM 工艺模块的计算机程序流程图

为了计算任意时刻任一点的杂质浓度，SUPREM 程序中考虑了杂质的再分布和每一工艺过程中发生的其他过程。每一节点的值是根据模拟区域边界的初始条件或前一步骤中每一节点的已知值来计算得到的。然后将被模拟的工艺过程划分为小的时间间隔 Δt，在每一时间间隔内发生的变化通过解相关的方程得到。例如，在最简单的描述一维扩散的 Fick 第二定律中（等式(2.5.7)）

$$\frac{\partial C(x)}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{2.9.1}$$

浓度关于空间的导数可由相邻节点间的浓度差代替；关于时间的导数由时间间隔 Δt 表示

$$\Delta C(x) = D\frac{\Delta^2 C}{\Delta x^2}\Delta t \tag{2.9.2}$$

那么，在 $t+\Delta t$ 时刻某一个节点的浓度由同一节点和相邻节点在 t 时刻的值决定

$$C(x, t + \Delta t) = C(x, t) + \Delta C(x) \tag{2.9.3}$$

在上面的例子中，计算二阶导数值要用到的不仅仅是最近邻节点的浓度值[$C(x-\Delta x, t)$ 和 $C(x+\Delta x, t)$]，才能确定浓度分布曲线的曲率。在数值模拟器中包括更复杂的扩散过程的模型，不需要将扩散系数近似为常数；扩散系数 D 可以是浓度 C 的函数，如等式(2.5.15)。

淀积、刻蚀、杂质淀积、氧化和向外扩散等过程发生在暴露的表面；离子注入发生在靠近表

面处;扩散在整个结构中发生。整个结构由单晶硅、多晶硅、二氧化硅、氮化硅、氮氧化硅、铝和光刻胶区域组成。这些材料层可被淀积或去除,在高温和应力下氧化硅可以流动。因此,在每一工艺过程中,必须考虑几种不同的物理过程。例如在氧化过程中,我们必须考虑杂质在衬底中的扩散,杂质原子从衬底向正在生长着的氧化物中的迁移,生长着的氧化物可能的黏滞性流动,以及氧化物生长的过程。在两种或多种材料的交汇处有多重节点,每一种材料在该处都有一个节点,所以可以表示出参数的多重值。在表示环境气体浓度的边界时,还会用到额外的节点。坐标系一般固定在衬底,而不是硅片的表面,以避免在氧化或刻蚀过程中表面位置变化引发的问题。

扩散 硅中典型的杂质原子包括硼、磷、砷和锑。在惰性气体或氧化气氛下,加热可引起杂质原子的扩散。在后一种情况,杂质原子和表面的硅都会相对于固定的衬底移动。正如我们在2.5节中看到的,扩散系数与杂质浓度和氧化或注入引起的点缺陷(硅间隙原子和空位)有关。我们得不到与浓度相关的扩散系数的总的解析解,但用数值模拟技术则很容易处理。氧化过程会从氧化界面向下面的硅衬底注入过剩的硅原子(自间隙原子)。这些点缺陷可以扩散很长的距离,从而影响到附近杂质原子的扩散。与之类似,离子注入引起的损伤可产生过剩的间隙原子和空位,并对附近杂质原子的扩散造成很大的影响。为了获得足够的精度,点缺陷对扩散的影响必须包含在模型中。内建电场对杂质扩散的影响也应包含在模型中;该电场把存在的不同杂质原子的扩散方程耦合在一起。

包含点缺陷效应的扩散系数模型为

$$D=\left[D^{0}+D^{+}\left(\frac{n_i}{n}\right)+D^{-}\left(\frac{n}{n_i}\right)+D^{=}\left(\frac{n}{n_i}\right)^{2}\right]F_{IV} \qquad (2.9.4)$$

D 的各种分量的含义在2.5节中讨论过;对每一杂质元素,不一定包含所有的分量。其中 F_{IV} 表示考虑了非平衡的点缺陷浓度引起的扩散系数的增加或减少的数量,与空位或间隙扩散机制造成的扩散部分有关。当杂质浓度较高时,电激活杂质浓度可小于总的客观存在的杂质浓度,这一效应及杂质在 Si-SiO_2 界面的分凝效应也包括在内。

氧化 氧化过程可用 Deal-Grove 模型来模拟,在时间 Δt 内生长的氧化物厚度的增量等于速率(等式(2.3.5)至(2.3.8))乘以时间增量。因为现在采用数值技术,不要求有精确形式的解析解,基本的 Deal-Grove 模型的速率表达式可用一项表示薄氧化物的初始氧化速率来修正,这一项将随着氧化物厚度的增加(等式(2.3.11))指数衰减[8]

$$\Delta x_{ox}=\left[\frac{B}{A+2x_{ox}}+K\exp\left(-\frac{x_{ox}}{L_o}\right)\right]\times\Delta t \qquad (2.9.5)$$

不同取向的单晶硅或多晶硅可采用不同的参数值。在二维模型中,氧化区域和受保护不被氧化的区域(例如被氮化硅层保护)之间的渐变区域用氧化剂的横向输运来确定。最后形成的氧化区域的形状示于图2.42。

离子注入 离子注入既可以用2.5节中的解析模型,也可以用蒙特卡洛方法来模拟。前者,Gauss 分布可以用投影标准偏差 ΔR_p 在表面和峰值浓度区间内(而不是采用峰值浓度以下的区域)的不同值来修正,也可以包含高阶偏差来修正。在晶格中沿晶向对离子的沟道效应进行建模方面也已具有一定的能力。在考虑横向偏差(y 方向)和纵向偏差(x 方向)时也包含

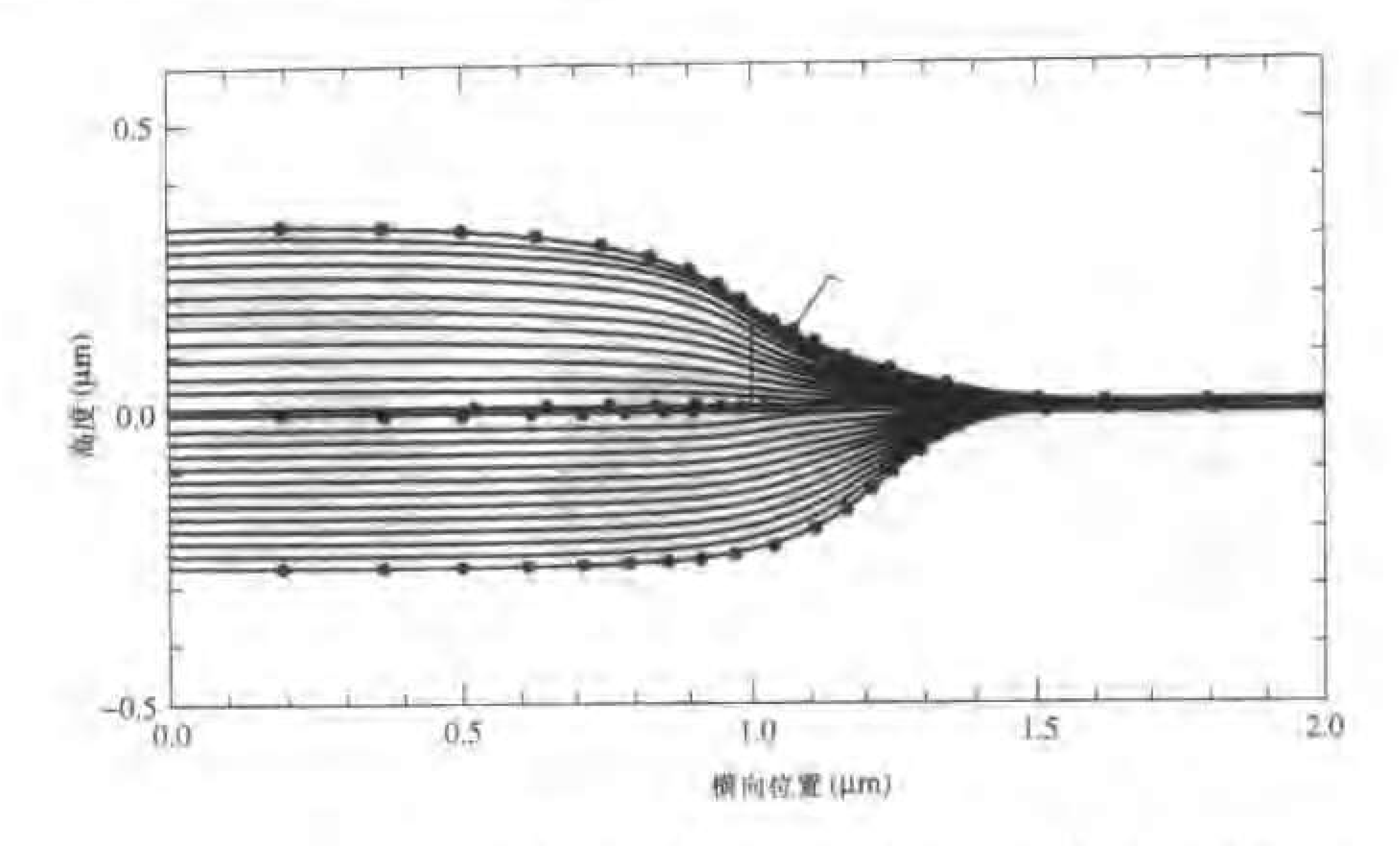


图2.42 在靠近氧化区和受氮化硅保护不被氧化区域的边界处，由氧化剂的横向输运形成的氧化物的二维图形（K. M. Cham 等，*Computer – Aided Design and VLSI Device Development*（Kluwer Academic Pubilishers，波士顿 1986）[21]）

了二维效应，所以注入的分布在 $y > |a|$ 时可写为

$$C(x, y) = C_p \exp\left[-\frac{x - R_p}{2\Delta R_p}\right]\left[\frac{1}{2}\operatorname{erfc}\left(\frac{y - a}{\sqrt{2}\Delta R_p}\right)\right] \tag{2.9.6}$$

式中，C_p 是无掩膜区的峰值浓度，a 是掩膜开口的半宽度。

另外，也可以用蒙特卡洛方法建模。考虑很多单个粒子的路径，直至对它们进行求和产生总的粒子分布图像。如果注入粒子必须通过不同材料组成的多层结构时，蒙特卡洛方法尤其有用。每一注入离子在每一次散射后可求出其能量的损失和方向的改变，当它进入下一层材料时这些量是已知的。当它的能量接近零时，离子达到了它最后稳定的位置，当计算了大量的离子轨迹后，统计每一深度间隔内停留的离子数就可得到它的分布。原子核和电子能量损失机制一般都需考虑。因为要重复计算大量的离子才能确定分布，蒙特卡洛方法很费时间，只有在更有效的解析模型不能给出足够的精度时才使用。为了减少计算时间，组合的参数有时简化为一个变量，将这一变量算出的解存于一个表中，对于给定的变量值可在表中找到相应的解，而不是每次需要时都计算一遍。

结合合适的模型，可对离子注入和退火后靠近复杂过渡区的杂质浓度进行二维建模。图2.43给出了器件与隔离区之间的过渡区的杂质分布等浓度线。

淀积与刻蚀 在高温淀积过程（例如外延）中，已存在于结构中的杂质会扩散。因此，除了要模拟所增加的这一层结构外，还必须计算所有杂质粒子的扩散。对于低温淀积，可忽略扩散以节约计算时间。刻蚀也可被详细定义。在基本的模型中，仅将材料去除至给定的深度，并简单地假定剩余材料的边界形貌。目前已研发出了独立的、复杂的模型来描述刻蚀边缘的形状（也包括淀积在不规则表面和沟槽内的淀积层形状）。

应力 计算应力很重要。如果在工艺过程（例如热氧化）中应力变得太大，就会产生晶格

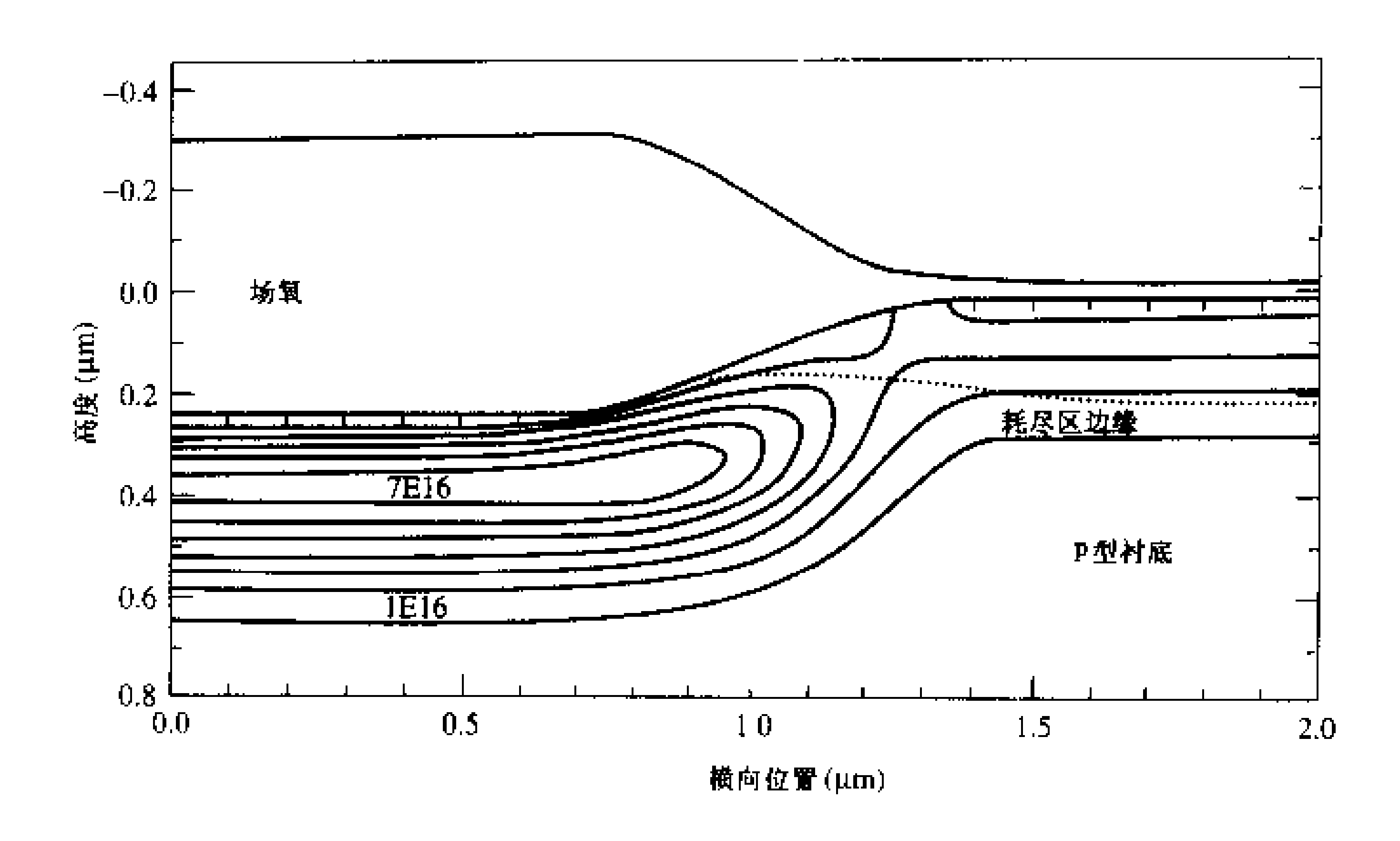


图 2.43 LOCOS 氧化物隔离区边界附近的杂质分布的等浓度线,与图 2.42 相似(Cham 等[22])

缺陷。例如,在 LOCOS 氧化过程(见 2.6 节)中,如果在硅与氮化硅之间用于应力释放的氧化硅层的厚度不够,将会在硅中产生大量的应力。这些应力会产生晶体线缺陷(位错),假如有位错线穿过 pn 结,该 pn 结的电学特性会变坏(由于这些应力会引起缺陷,因此要增加一层应力释放氧化层,尽管这样氧化区会向横向延伸,使得精确的尺寸控制变得困难)。应力的计算要用到杨氏模量、泊松比和每种材料的线性热膨胀系数。由淀积工艺引起的本征应力也可包含在内。

氧化过程中计算得到的应力可用来修正氧化反应速率常数和氧化剂在已形成的氧化硅中的扩散系数。在高温下,SiO_2可以发生回流,这种流动也可包含在模型中。在考虑部分用掩蔽层覆盖的表面氧化时回流尤其重要,因为回流能够改变过渡区的形状。模型也能预测样品冷却后留在薄膜中的残余应力。由于 Si、SiO_2和其他材料的热膨胀系数不同,即使在氧化温度下不存在应力,当样品冷却时应力也会增加。

表面 在研究硅中杂质原子浓度和它们在工艺过程中的变化时,可以使用纯粹的数值模拟。在某一点的变化量通常主要由邻近的点在前一个时段的值决定。然而,靠近表面处,某一点的变化主要由从远处到达的粒子数决定。例如,淀积和气相刻蚀都受到气相条件下一个平均自由程内粒子数目的影响,这个距离通常大于器件的特征尺寸。在某些情况下,可采用蒙特卡洛方法。单个粒子以任意速度和方向到达硅片,它们的轨迹和作用可以用已知影响它们运动规律的物理和化学定律来模拟。考虑了大量的粒子后,就可以决定其平均的行为;但是,大量的蒙特卡洛模拟所需的计算资源限制了将其用于对特定细节的研究。

对于一般的应用,准解析解通常更引人注目。例如,求解某一点的淀积或刻蚀速率,可通过对远处气态源(例如物理气相淀积源或 CVD 中的化学粒子)直接到达该点和到达其他相邻表面后反射到该点的粒子(离子、原子团等)作用进行积分得到。表面其他点随时间的变化可采用类似的方法进行计算

后处理 一旦模拟完成,通常将数据发送到另外的后处理程序中进行绘图和数据处理,以

帮助理解模拟的结果。

例题 工艺模拟:一维硼扩散[23]

利用模拟程序 SUPREM 研究硅片中(1)离子注入硼(B)并(2)经过随后几个高温工艺过程,求工艺完成后硼的分布。

解: SUPREM 的输入数据包括以下几行计算机代码。首先要给出标识分析过程的标题。

TITLE: BORON IMPLANT REDISTRIBUTION

其次,要描述硅衬底,包括杂质类型、浓度和晶向。

SUBS ELEM = P, CONC = 2E15, ORNT = 100

在这个例子中,衬底的掺磷(P)浓度为 $N_d = 2\times10^{15}\mathrm{cm}^{-3}$,属于(100)晶向。

然后,描述与垂直方向网格间距有关的参数。硅中的垂直方向分为两部分,一部分为非常接近表面的高精度区,一部分为远离表面的低精度区。网格间距、高精度区的深度和总的计算深度由使用者来确定。一般在垂直方向上取 350 到 400 网格点。那些没有指明为高精度区内的点会自动地归属于低精度区。

GRID DYSI = 0.005, DPTH = 1.5, YMAX = 2.5

这一行定义了高精度区的网格间距 ΔY 为 0.005μm,深度为 1.5μm(因此,在这一区域用了 300 个点)。总的模拟区域延伸到硅表面下 2.5μm 深处,因此低精度区域可以有 50 到 100 个网格点。

此时可以定义打印和绘制计算结果的详细输出格式。另外,也可定义在基本的 SUPREM 程序中没有包括的工艺模型。如果没有另外定义的话,计算中就采用内嵌的模型。

输入这些初始参数后,要说明实际的注入、扩散和氧化工艺条件。一行语句定义一步工艺。

STEP TYPE = IMPL, ELEM = B, DOSE = 3.2E13, AKEV = 380

在这个例子中,注入硼的剂量为 $3.2\times10^{13}\mathrm{cm}^{-2}$,注入能量为 380keV。接着定义随后的热处理工艺。其中一个可能的步骤如下:

STEP TYPE = OXID, TEMP = 1000, TIME = 30, MODL = DRY1

这里定义的氧化温度为 1000℃,时间为 30 分钟。氧化动力学已由以前输入的模型 DRY1 描述。它是指硅片插入氧化炉时采用的是稀释的氧环境。也可定义其他氧化和惰性气氛热处理环境。

模拟工作完成后,打印出程序计算得到的参数,并画出最终的杂质分布。在本例中,有关的参数是硼在硅、氧化硅中的浓度分布和表面的浓度。经过上面描述的两个工艺步骤并接着经历了类似于制造 MOS 晶体管时用到的四个热处理工艺后,使用 SUPREM 对硼杂质分布的推算结果示于图 2.44 中。

2.9.4 器件模拟

工艺模拟是电子器件全部模拟工作中必不可少的第一步。它可以利用一维、二维或三维空间,告诉我们详细的物理结构和杂质形貌。这些结果然后被用做对器件电学特性进行详细

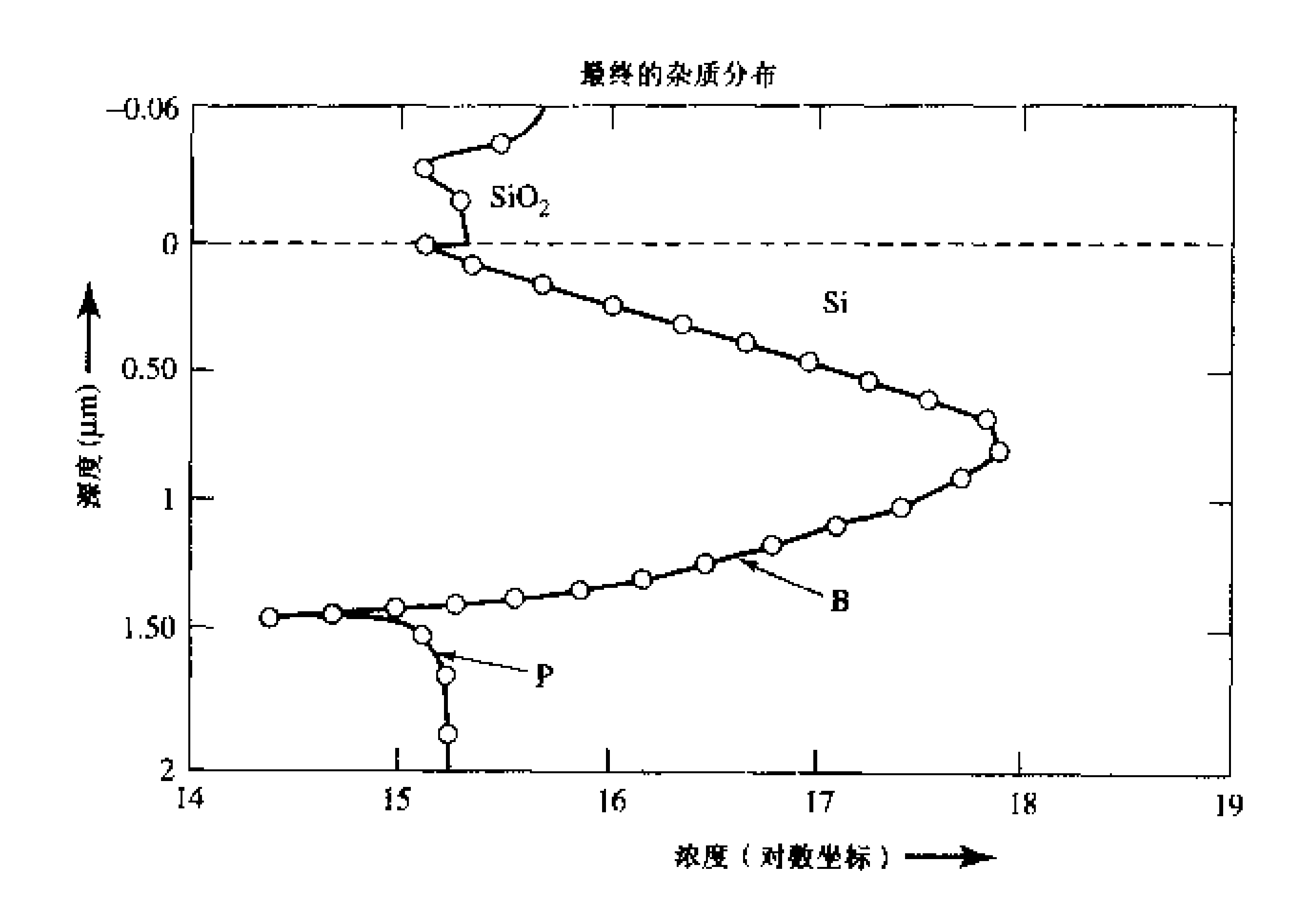


图 2.44 工艺过程结束后,使用 SUPREM 计算得到的硼和磷的分布

模拟的起点。工艺模拟程序的输出通常存在计算机文件中,可以直接用于器件模拟程序的输入。这种工艺与器件模拟的耦合极大地增强了从工艺流程到器件性能的预测能力。对工艺流程进行自动优化以得到所需要的器件特性是下一步必然要做的工作,但是难度要大得多

与工艺模拟一样,器件模拟也可以是一维的或多维的。一维模型可模拟更复杂的物理效应。由于现有计算机能力的限制,还不能对一些效应进行二维或三维模拟。器件模拟程序可提供关于稳态和瞬态条件下的器件特性。一般的模型基于对 Poisson 方程和电流连续性方程进行求解,并同时考虑电子和空穴的贡献。被模拟的物理结构可以直接在器件模拟器的输入文件中进行定义。这一结构可利用定义的尺寸和杂质分布的解析模型得到,或进行近似估算得到,或将工艺模拟器得到的输出文件用做器件模拟器的输入来预测出器件的结构。在第 5 章中学习了更多的器件概念后,我们将对器件模拟进行更详细的讨论。

2.9.5 模拟面临的挑战

随着器件尺寸的减小,在设计先进的工艺和器件时必须考虑更高阶的物理效应,这样工艺和器件模拟就变得更加重要了。然而,当器件变得更小时,描述器件特性和工艺的详细的物理模型变得越来越复杂,有些现象还没有被完全理解。尽管模型在不断发展,但是在进行先进的工艺和器件模拟时,缺乏合适的物理模型是面临的主要挑战。使用今天甚至昨天的物理模型来模拟未来的工艺、器件和电路并不能提供足够的预见性。根据不适当的物理模型进行精确的计算可能会误入歧途,尤其是当使用者越来越不了解模型中的细节时。

由于这些实际问题的限制,靠模拟不能代替对相关物理机制的理解。然而,与仅仅依靠实验相比,通过模拟能够对器件有更好的理解。由于它不像解析模型必须对有限的方程进行近似,因此数值模拟能提供比解析计算更多的信息。甚至当某些参量的精确值不知道时,仔细地进行模拟也可揭示出器件特性的变化趋势。另外,必须经常简化工艺模型以使计算能在合理的时间内完成。当对许多工艺组成的流程进行模拟时,每一步工艺的近似引入的小误差混合

在一起,最后的计算结果可能会造成相当大的误差。必须采用实验测量来证实某些关键点的计算结果,确定出模型参数,以及指出改进模型的方法。在理想情况下,模拟提供了将我们对物理机制的理解与实验结果结合在一起的工具。

2.10 器件:集成电路中的电阻

电阻是许多集成电路中简单而又重要的电子元件。采用本章已讨论的工艺过程,可有多种不同的方式制作电阻。在对扩散的IC电阻的电学特性进行全面的讨论后,我们再来描述电阻的制造工艺。

在第1章中提到,均匀导电的固体棒状材料的电阻由下面的等式给出

$$R = \frac{\rho L}{A} \tag{2.10.1}$$

式中,电阻率ρ为等式(1.2.7)中定义的电导率的倒数,这里可写为

$$\sigma = \frac{1}{\rho} = (q\mu_n n + q\mu_p p) \tag{2.10.2}$$

形成集成电路电阻的一个常用方法是在均匀掺杂硅衬底上的保护层SiO_2上开窗口,然后引入与硅片掺杂类型相反的杂质,如图2.45所示。在第4章中将会看到,在导电类型相反的两个区域的交界处会产生阻止电流流动的势垒。因此,如果在p型区的两个端点附近形成接触电极,加电压时在这一区域将有平行于表面的电流流过。我们不可能用等式(2.10.1)来计算这一电阻的阻值,因为它不是均匀的棒状材料。如图2.21所示,上述工艺形成的杂质浓度从靠近表面处的极大值向硅内部逐渐下降。为了计算这种情况下的电阻,考虑平行于表面的电导会有所帮助。

电导 考虑将p型杂质掺入到n型硅片中形成的一个p型电阻,如图2.45所示。在深度为x处,对于一平行于表面、厚度为dx的薄层p型区域,它的微分电导dG为(如图2.45b所示)

$$dG(x) = q\mu_p p(x)\frac{W}{L}dx \tag{2.10.3}$$

对从表面一直到底部的每一薄层电导相加,就可以计算出整个p型区域的电导G。在对薄层厚度取极限的情况下,这个和式变为积分

$$G = \frac{W}{L}\int_0^{x_j} q\mu_p p(x)\,dx \tag{2.10.4}$$

式中,x_j表示空穴浓度可忽略时的深度(接近$N_a = N_d$处)。

如果p型区域是由气相淀积和随后的推进扩散形成的,那么可以将杂质分布$N_a(x)$用Gauss分布近似(等式(2.5.13))。正如2.5节中介绍的,扩散进入到硅片中的单位面积总的杂质浓度N'由淀积过程的特性决定,而扩散区的特征长度$2\sqrt{Dt}$与推进扩散有关。在硅片内的任意深度处,将p型杂质原子浓度减去原来的n型硅片杂质浓度所得到的净掺杂浓度近似为该处的空穴浓度。实际上,扩散区大部分电流是在杂质浓度最高的区域内流过的,通常为$10^{18}cm^{-3}$量级或更大。因为硅片的初始杂质浓度约为$10^{15}cm^{-3}$,我们可以忽略背景掺杂,并假

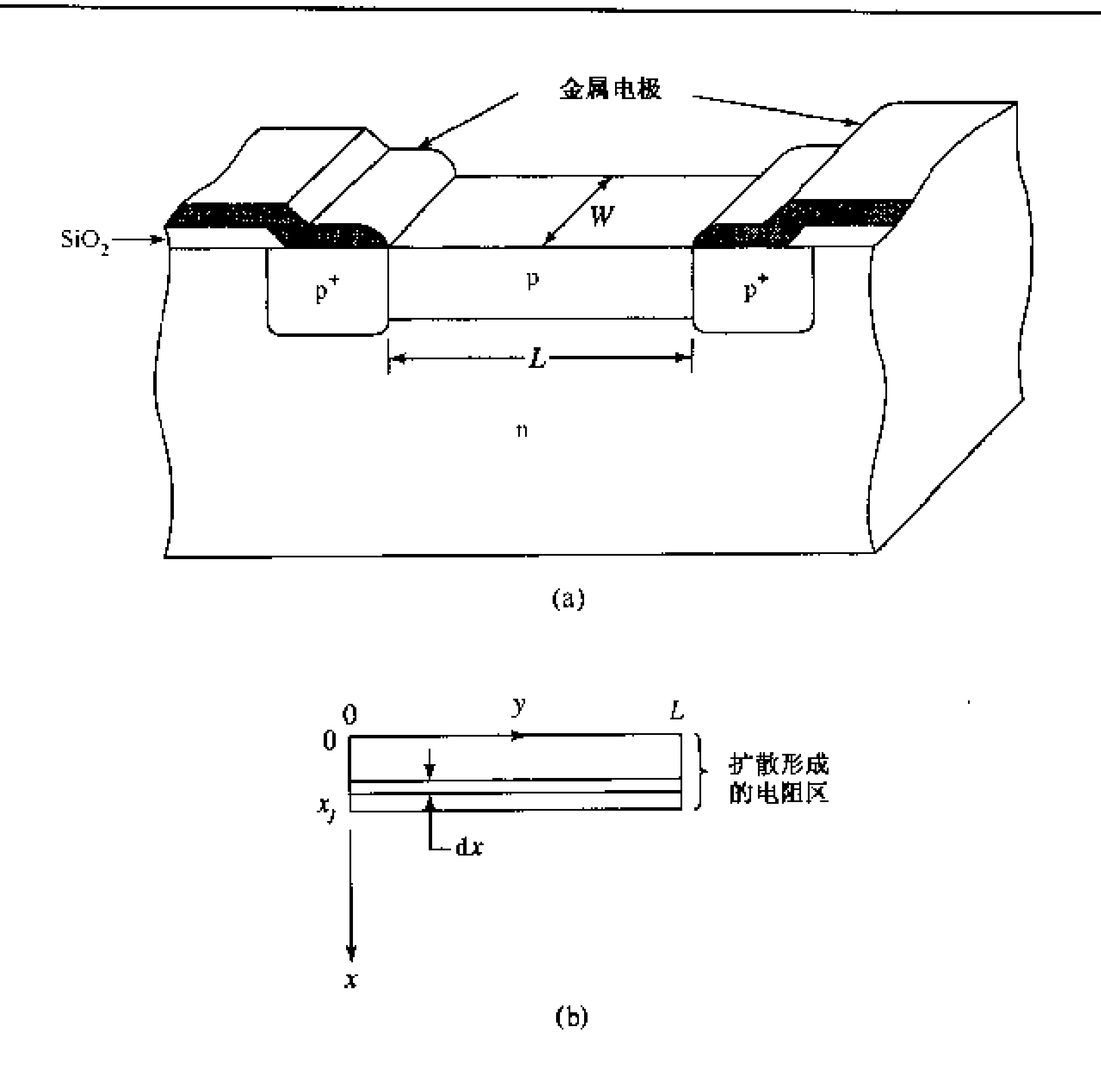


图2.45 (a)将受主杂质扩散进入到n型硅片上选择的区域中形成IC电阻。p^+区是重掺杂区域,以保证金属电极与p型电阻区间良好的接触。(b)由等式2.10.3给出的具有电导dG的电阻器中一薄层区域

定$p(x)=N_a(x)$[④]。将等式(2.5.13)代入到等式(2.10.4)中得到

$$G=\frac{qN_p^t}{\sqrt{\pi Dt}}\frac{W}{L}\int_0^{x_j}\mu_p\left[\exp\left(\frac{-x^2}{4Dt}\right)\right]dx \tag{2.10.5}$$

等式(2.10.5)中的迁移率μ_p是杂质浓度的函数(见图1.6),从而也是与表面距离的函数。因此,不能将它移到积分号的外面。

计算等式(2.10.5)最直接和最精确的方法是用计算机进行数值积分。另一种经常采用的有价值的方法是对图1.6中曲线的不同区域用解析表达式近似,然后用于等式(2.10.5)。两种方法在原来的假设限制下都具有很高的精度。然而,如2.5节所述,真实的扩散分布会偏离简单的Gauss分布。因为这个原因,只有对被研究的特定杂质离子的扩散分布进行了实验研究或者拥有如2.9节中讨论过的更精确的工艺模型时,才能保证详细的数值分析结果。

我们用迁移率的平均值$\bar{\mu}_p$可计算出电导的近似值。因为大部分电流是在杂质浓度接近于最大值处流过,选择对应于最大杂质浓度一半时的迁移率值为其平均值是合理的,而且与采用的简化杂质扩散分布一致。这样,电导的表达式(等式(2.10.4))可简化为

④ 这里假定空穴和受主原子具有相同的深度分布。在4.1节中我们将指出这一假设是不太精确的。在目前这种情况下,N_a和p之间小的差异可被忽略。

$$G = N'q\overline{\mu_p}\frac{W}{L} = g\frac{W}{L} \tag{2.10.6}$$

式中,$g \equiv N'q\overline{\mu_p}$ 是一方形电阻的电导($L = W$)。电导由平均迁移率 $\overline{\mu_p}$ 与单位表面积上总的杂质浓度 N'(等式(2.5.10))之积决定。因此,电阻 R 为

$$R = \frac{1}{G} = \frac{L}{W}\frac{1}{g} \tag{2.10.7}$$

在集成电路制造中,通常采用在同一块掩膜上定义出不同几何形状的方法同时制备出许多电阻。因为所有电阻在相同的扩散过程中形成,因此将电阻的大小分为两部分比较方便:L/W之比由掩膜尺寸决定,而 $1/g$ 由扩散工艺决定。

方块电阻 掩膜上的任意电阻图形可分为每边边长均为 W 的一系列方块(图 2.46)。图形中方块的数目恰好等于 L/W。所以,电阻值等于可划分的方块数目与通常用符号 $R_{\square}$(有时为 R_s)表示的参数 $1/g$ 之积,$R_{\square}$称为方块电阻。方块电阻的单位是欧姆,有时为方便也表示为每方块欧姆(Ω/□),以强调一个电阻值是由方块数和方块电阻相乘得到的。例如,一个长为 100μm、宽为 5μm 的电阻包含 20 个方块(20□)。如果采用的扩散工艺使扩散层的方块电阻为 200Ω/□,则电阻的阻值为 20□ ×200Ω/□ =4.0kΩ。

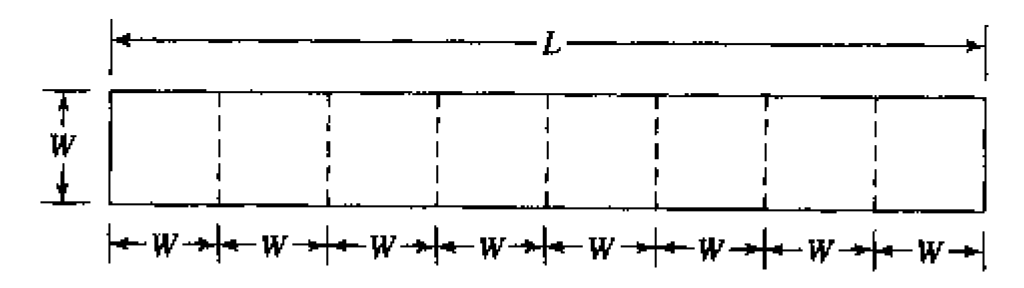


图 2.46 描述一个电阻表面尺寸的方块数由比例 L/W 给出

方块电阻阻值可控制在千欧量级,阻值越高需要包含的方块数越多。因为图形的宽度由掩膜的尺寸和刻蚀极窄线条的能力决定,对高值电阻而言,电阻的长度可变得很长以获得所需的方块数。高值电阻需要大的面积,这实际上成为集成电路中的一个制约因素,设计的电路通常避免使用高值电阻。在电路允许的情况下,通常采用晶体管来代替高值电阻,因为晶体管"廉价"(即占用较少的表面积)。当必须使用含有大量方块数的电阻时,通常将其设计为如图 2.47 所示的弯曲形状,以减小占用的面积。电流流过这种图形的拐角方块时是不均匀的。拐角方块的电阻可估计为直线方块电阻阻值的 65%。

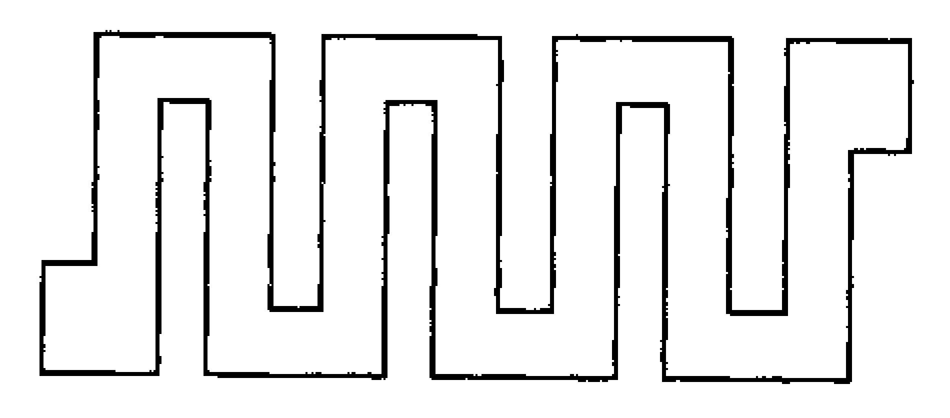


图 2.47 当必须设计长的高值电阻时可采用折叠图形

对于大面积电阻,L 和 W 可简单地由掩膜尺寸决定。然而,对于非常窄的电阻(W 小),有

效宽度 W_{eff} 可能与掩膜尺寸有很大不同,因为杂质在氧化硅下面除了垂直扩散之外,同样也会横向扩散(图 2.48)。如果 W 远大于扩散深度 x_j,可忽略这一效应。然而,为了使给定的方块数目占据最小的面积,通常选取的 W 值小至掩膜公差许可的尺寸。在这种情况下,x_j 与 W 相比已不能被忽略,必须使用有效值 W_{eff} 来计算考虑横向扩散后的电阻宽度。

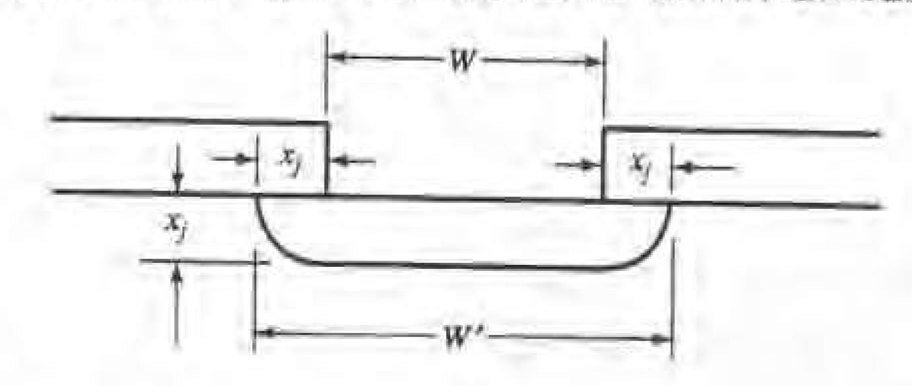


图 2.48 横向扩散使电阻尺寸不同于标称的掩膜尺寸

例题 扩散电阻

在某集成电路中设计的一个 p 型电阻是由两个高电导的 p 型接触区连接的电阻条,两接触区间的电阻条长 4μm、宽 1μm。电阻条的结深为 1μm,需要设计的电阻值为 1kΩ。

确定方块电阻和满足给定指标需要的平均电阻率(本例题忽略横向扩散)。

解:电阻图形的方块数为

$$L/W = 4/1 = 4$$

因此,方块电阻 R 为

$$R = 1000/4 = 250\Omega/\square$$

根据等式(2.10.6)和关系式 $R_{\square} = (1/g)$,我们可计算出单位面积需要的杂质浓度 N'_a

$$N'_a = (q\bar{\mu}_p R_{\square})^{-1} \text{ 掺杂原子 }/\text{cm}^2$$

平均的体杂质浓度 $\overline{N}_a$ 与面杂质浓度 N'_a 之间的关系为

$$\overline{N_a} = \frac{N'_a}{x_j}$$

则 p 型扩散电阻的平均电阻率为(等式(2.10.2))

$$\bar{\rho} = (q\overline{\mu_p}\overline{N_a})^{-1} = \left[\frac{q\overline{\mu_p}N'_a}{x_j}\right]^{-1} = R_{\square}x_j = 250 \times 1 \times 10^{-4} = 0.025\ \Omega\cdot\text{cm}$$

电阻值的精度 在离开扩散电阻的主题之前,还有一点需要强调。在等式(2.10.7)中已经看到,电阻是两个因子的函数,其中 L/W 由采用的光刻控制,方块电阻 $1/g$ 或 $R_{\square}$ 取决于杂质的淀积和再分布。方块电阻的控制通常是精密电阻设计中的最重要的限制因素。尽管在工艺中得到的 $R_{\square}$ 值可能会有些变化,但是它在一个典型的芯片面积上几乎为常数。这样,在 IC 中两个相邻的电阻应该具有相同的方块电阻值,两者的电阻值之比应由它们的相对的尺寸决定。采用合适的工艺,这些尺寸可以精确地控制,从而相邻电阻阻值之比要比电阻阻值本身的精度更高。因此,在设计集成电路时,往往使它的关键特性取决于两个电阻阻值之比,而不是依赖于电路中某一指定电阻的阻值。

因为与其他平面工艺兼容,扩散电阻在硅集成电路中获得了广泛应用。它们通常与其他

电路元件同时形成，因此不增加制作成本。但是，不采用扩散工艺，在集成电路中通过图形化与下面衬底形成pn结隔离的外延层也可能制作出电阻。典型的双极晶体管集成电路就是在这种外延层上制作的。因为在制作电路的硅材料中外延层的电阻率是最高的，因此这种方法得到的方块电阻（约1000Ω/□）比在p型衬底上通过正常扩散得到的方块电阻（约200Ω/□）大好几倍。对于更高的方块电阻，有时采用双扩散结构，即在p型区上形成n型区以减小扩散电阻的垂直尺寸。这种电阻称做夹断电阻，可使方块电阻增加约40至50倍。制作夹断电阻显然更复杂，并且方块电阻的重复性一般较差。

如果某一特殊电路要求的电阻值很大或很精确，可采用另一种方法。不通过对硅片掺杂来形成电阻，而是在覆盖大部分电路的绝缘二氧化硅上淀积一层电阻膜。然后，对这一层膜用掩膜图形化并用金属线与电路的其他部分相连。使用淀积的电阻给电路设计提供了很大的灵活性。有时，这些电阻可用IC工艺中已有的一层材料形成。例如，使用淀积的多晶硅，它已是硅栅MOS集成电路制造工艺的一部分。如果采用的电阻层不是给定的IC工艺中的一部分，必须仔细考虑对这一层电阻材料的淀积和图形化所需增加的成本。

在IC电阻的讨论中，我们已经看到工艺带来的特殊限制影响到了电路设计。在随后几章讨论其他IC器件时，我们将会不断看到这种情况。

小结

硅材料对电子学的无与伦比的重要性来自其优越的材料特性和已经取得的对加工工艺极好的可控性。在硅上能够制作出高度绝缘的氧化物，它不仅具有极好的重复性，同时在器件与氧化物之间具有很好的界面可控性，这些特性都是独一无二的。用于制作硅集成电路的平面工艺是精确制造小尺寸器件的基础。它使得同时制作许多器件成为可能，因此是获得均匀、可靠和经济的IC生产的关键。

IC制造的原材料——大直径的单晶硅通常用直拉法制造。在氧含量必须很低的情况下则采用区熔法。在平面工艺开始前，需将单晶硅锭切割成硅晶圆片。由硅氧化生成的SiO_2，既可在接近1000℃的干氧气氛下，也可在类似温度的湿氧气氛下（速率快得多）进行。IC中复杂的图形需用光刻技术对光敏聚合物（光刻胶）薄膜图形化得到。采用图形化后的光刻胶作为掩蔽层，能选择性地去除IC工艺中的氧化层和其他材料。杂质能改变硅的电导率（包括电导率的大小和导电类型），可以在用光刻胶限定的图形，或光刻胶在其他材料上所限定的图形中将杂质引入到硅中。离子注入和气相淀积都可以在IC所需的区域内引入杂质。随后的杂质原子扩散进入硅片的过程可用最简单的Fick定律来描述，这是一个偏微分方程，针对具体的集成电路的扩散过程对应有两个解析解（Gauss分布和余误差分布）。然而，Fick定律不能解释一些更复杂的扩散效应，对于更精确的IC工艺设计需要采用计算机模拟来预测扩散过程。在单晶硅衬底上制作单晶硅（外延），淀积多晶硅和绝缘层所用到的化学气相淀积是一种重要的IC工艺。

集成电路中器件的互连是一个重要工艺步骤，对互连使用的导电材料有着严格的要求。铝容易发生电迁移，随着器件和导线尺寸的减小，会变得更加严重。由于铜的电阻率低且不易发生电迁移，已在高性能电路中使用。

由周期表中Ⅲ族和Ⅴ族元素组成的化合物半导体通常用于特殊器件。某些化合物半导体

在光学器件中特别有用,因为具有直接带隙,当电子和空穴复合时可产生有效的光发射。另外,由于它们的迁移率很高,还可以用在高频器件中。对于某些半导体材料具有相同的晶格常数但不同的带隙,则很容易将它们组合在一起制造出有用的器件。

在无法进行昂贵的工艺实验的情况下,采用复杂的计算机程序进行工艺和器件模拟,为工程师研究和设计物理结构、杂质分布和预测器件性能提供了一个必需的工具。功能强大的数值技术允许模型中包含越来越多的物理效应。当器件更小时,这些效应变得更加重要。

扩散电阻在集成电路中应用广泛,通常采用平面工艺,通过限定载流子的电流通道图形来实现。电阻值可用方块电阻计算,测量单位为每方块欧姆(Ω/□),用符号 $R_□$ 表示。方块电阻表示通过正方形两对边的电阻;因此扩散电阻的阻值可用组成电阻图形的方块数与 $R_□$ 相乘得到。

参考文献

1. *Dilbert* by Scott Adams, United Features Syndicate, 7/15/1997.
2. R. Chau, et al., Tech. Digest Int'l. Electron Devices Mtg. (San Francisco, December 10–13, 2000), pp. 45–48; 2001 Silicon Nanoelectronics Workshop, Kyoto, June 10, 2001.
3. International Technology Roadmap for Semiconductors: http://public.itrs.net; 2001 update: http://public.itrs.net/Files/2001ITRS/Home.htm.
4. W. C. O'Mara, R. B. Herring, and L. P. Hunt, eds., *Handbook of Semiconductor Silicon Technology*, (Noyes Publications, Park Ridge NJ, 1990).
5. B. E. Deal and A. S. Grove, J. Appl. Phys. **36**, 3770 (1965).
6. R. R. Razouk, L. N. Lie, and B. E. Deal, *J Electrochem. Soc.* **128**, 2214 (October 1981).
7. D. Hess and B. E. Deal, *J. Electrochem. Soc.* **124**, 735 (1977); J. J. Barnes, J. M. DeBlasi, and B. E. Deal, *J. Electrochem. Soc.* **126**, 1779 (1979); B. E. Deal, J. Electrochem. Soc. **125**, 576 (1978); F. Shimura and H. R. Huff in *VLSI Handbook* (ed. N. G. Einspruch), Academic Press, Orlando, FL (1985), Chapter 15.
8. H. Z. Massoud, J. D. Plummer, and E. A. Irene, *J. Electrochem. Soc.* **132**, 2685 (1985).
9. C. Ho, J. D. Plummer, and J. D. Meindl, *J. Electrochem. Soc.* **125**, 665 (April 1978).
10. M. C. H. M. Wouters, H. M. Eijkman, and L. J. van Ruyven, Philips Research Rep. **31**, 278 (1976).
11. J. F. Gibbons, W. S. Johnson, and S. W. Mylroie, *Projected Range Statistics*, 2nd Ed. Dowden, Hutchinson, and Ross, New York (1975).
12. Research Triangle Institute, *Integrated Silicon Device Technology, Vol. IV, Diffusion* (ASD-TDR-63-316). Research Triangle Institute, Durham N. C. (1964). Reprinted by permission of the publisher
13. F. A. Trumbore, *Bell System Tech. J.* **39**, 205 (1960); G. Masetti, D. Nobili, and S. Solmi, *Semiconductor Silicon 1977*, Electrochemical Society (1977), p. 648; A. Armigliato, D. Nobili, P. Ostoja, M. Servidori, and S. Solmi, *Semiconductor Silicon 1977*, Electrochemical Society (1977), p. 638; D. Nobili, A. Carabelas, G. Celotti, and S Solmi, *J. Electrochem. Soc.* **130**, 922 (April 1983); R. A. Craven, *Semiconductor Silicon 1981*, Electrochemical Society (1981), p. 254.
14. R. B. Fair, *Semiconductor Silicon 1977*, Electrochemical Society (1977), p. 968.
15. E. S. Meieran and T. I. Kamins, Solid-State Electr. **16**, 545 (1973).
16. B. Swaminathan, *Doctoral dissertation*, Department of Elect. Eng., Stanford University (April 1983).
17. J. A. Appels, E. Kooi, M. M. Paffen, J. J. H. Schatorje, and W. H. C. G. Verkuylen, Philips Research Rep. **25**, 118 (1970).
18. P. P. Merchant, *Hewlett-Packard Journal*, **33**, 28 (August 1982).
19. P. Vande Voorde, S.-Y. Oh, and R. W. Dutton, private communication.
20. D. A. Antoniadis and R. W. Dutton, *IEEE J. Solid-State Circuits* **SC-14**, 412 (April 1979); R. W. Dutton and Z. Yu, *Technology CAD: Computer Simulation of IC Processes and Devices*, Kluwer Academic Publishers, Boston (1993).
21. K. M. Cham, S. Y. Oh, D. Chin, and J. L. Moll, *Computer-Aided Design and VLSI Device Development*, Kluwer Academic Publishers, Boston, 1986, p. 60.
22. Cham, et al., p. 221.
23. R. D. Rung, "Silicon IC Technology Series: Computer Simulation of Silicon Processing," Videotape No. 90862, Hewlett-Packard Co. (1979).
24. F. M. Smits, *Bell System Tech. J.* **37**, 711 (1958).

参考书

S. M. SZE, *Semiconductor Devices: Physics and Technology*, second edition, Wiley, New York (2002).

J. D. PLUMMER, *Silicon VLSI Technology: Fundamentals, Practice, and Modeling*, Prentice Hall, 1999.

S. WOLF and R. N. TAUBER, *Silicon Processing for the VLSI Era: Vols. I–3*, Lattice Press, Sunset Beach CA (1990–2000).

S. K. GHANDHI, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, second edition, Wiley-Interscience, New York (1994).

W. E BEADLE, J. C. C. TSAI, and R. D. PLUMMER (Editors), *Quick Reference Manual for Silicon Integrated-Circuit Technology*, Wiley-Interscience, New York (1985).

S. M. SZE *Physics of Semiconductor Devices*, second edition, Wiley, New York (1981).

S. M. SZE (Editor), *VLSI Technology*, McGraw-Hill, New York (1983).

R. A Colclasser, *Microelectronics, Processing, and Device Design*, Wiley, New York (1980)

习题

2.1* 用直拉法生长晶体硅。晶体生长之前，在坩埚中10kg的熔融硅中加入了1mg的磷

(a)在开始晶体生长时，固体中初始的杂质浓度是多少？

(b)在5kg溶化的硅结晶后，硅晶体表面的杂质浓度是多少？

[磷在硅中的分凝系数$\frac{C_{固}}{C_{液}}$为0.3。]

2.2 将直拉法生长的硅片在高温氮气中加热，使靠近硅片表面区中的氧蒸发。然后在低温下加热引起残余的氧集结成团 解释这一工艺过程是如何和为什么提高随后制作的器件的电学特性的。

2.3 (111)晶向的硅片在IC工艺中多次氧化。计算按下面顺序进行的每一次工艺后二氧化硅的总的厚度。

(a)在1100℃的干氧和HCl(加入足够的HCl比在纯的O_2中提高氧化速率约10%)中60分钟。

(b)在1000℃的水蒸气(1个大气压)中2小时。

(c)在1100℃的干氧中6小时。

2.4* 硅片上覆盖有200nm厚的二氧化硅。在1200℃的干氧中再生长100nm厚的二氧化硅需要多长时间？

2.5 (a)在1000℃、1个大气压下的水蒸气中生长1μm的二氧化硅需要多长时间？[问题(a)至(d)需考虑(111)晶向。]

(b)在1000℃、10个大气压下的水蒸气中生长1μm的二氧化硅需要多长时间？

(c)在800℃、1个大气压下的水蒸气中生长1μm的二氧化硅需要多长时间？

(d)在800℃、10个大气压下的水蒸气中生长1μm的二氧化硅需要多长时间？

[这个问题表明：在低温(800℃)下采用高压的方法也能生长出厚氧化硅(1μm)。]

2.6 利用等式(2.3.5)推导出等式(2.3.6)。

2.7* 在某一LOCOS工艺中(已在2.6节中介绍)，淀积50nm厚的氮化硅层并图形化后，在水蒸气(一个大气压，1000℃)中进行了8个小时的氧化。氧化后，去除氮化硅层，暴露出硅表面。请问硅表面与生长的氧化硅表面的距离是多少(8小时的水汽氧化中，约有24nm的氮化硅转变为氧化硅)？计算(100)晶向的硅片

2.8 在LOCOS工艺中，请问每氧化1nm的氮化硅，在氮化硅层表面上生成的氧化硅的厚度是多少？

2.9* 在硅衬底上刻蚀出1μm宽和几个μm深的垂直槽。沟槽的表面是裸硅，但是在平的硅表面上覆盖有一薄氮化硅层作为氧化的掩蔽膜(图P2.9)。然后，硅片在1100℃、1个大气压下的水蒸气中氧化，用氧化硅填充沟槽。

(a)当沟槽完全填满时，二氧化硅条的宽度是多少?

(b)用二氧化硅填满沟槽所需的时间是多少?

[提示：注意，(a)中 x 单位宽度的氧化硅条是由($x-1$)单位的硅形成的. 对于(b)，利用等式(2.3.6)，取 $\tau=0$ 和图 2.7a 和 b 中(100)硅的数据。]

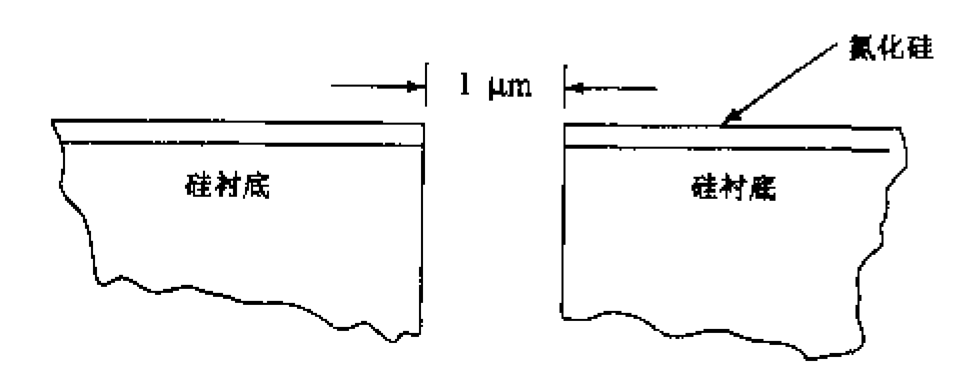


图 P2.9

2.10[1] 为了形成晶体管的接触区，注入到硅片中磷的剂量为 $3\times10^{16}\text{cm}^{-2}$，注入能量为 50keV($R_p=63\text{nm}$，$\Delta R_p=27\text{nm}$)。

(a)如果现在对硅片进行氧化，浓度增强氧化效应是否需要考虑?

(b)氧化前，磷在 1000℃扩散 60 分钟. 这时浓度增强氧化效应是否重要?

(c)如果磷的注入能量是 150keV($R_p=180\text{nm}$，$\Delta R_p=64\text{nm}$)重新考虑问题(a)和(b)。

2.11 高浓度的硼注入到 n 型硅表面后，在一部分硅片表面覆盖一层含有高浓度晶体缺陷的多晶硅. 然后对硅片进行氧化，发现在氧化过程中没有被多晶硅覆盖的区域的结深要大得多。请解释这一结果

2.12 玻璃的线性膨胀系数为 $9\times10^{-6}/℃$。假定两次光刻掩膜版的温度差 1℃而硅片温度保持不变;在这种条件下加工直径为 20cm 的硅片，且中心完全对准。请问在硅片边缘最小的对准误差是多少?

2.13 在 975℃时气相淀积 30 分钟将磷加入到硅片中. 确定以下衬底对应的结深:

(a)0.3Ω·cm 的 p 型硅衬底。

(b)20Ω·cm 的 p 型硅衬底。

假定在 975℃时，磷的扩散系数是 $10^{-13}\text{cm}^2\text{s}^{-1}$，固溶度为 10^{21}cm^{-3}。

2.14[†] 面密度为 10^{15}cm^{-2} 的硼原子从用载气稀释的 BCl_3 气态源进入硅片. 初始硅片的均匀施主浓度为 $5\times10^{15}\text{cm}^{-3}$。随后的推进扩散在 1100℃氮气中进行。需要的结深为 2μm。

(a)推进扩散需要持续多长时间?

(b)在 $\log N_a$ 和 N_a 与 x 的关系图中给出硼的浓度分布.

(c)需要设计一个 500Ω 的电阻. 最小的掩膜尺寸为 4μm. 掩膜上电阻的长度应为多少? 使用的硅表面积是多少? 对这个问题得到精确完整形式的解是不可能的，但是(b)中的图可指出求解的方法。给出近似解

(d)定性地描述推进扩散时间增加一倍的结果. 注意，净的受主杂质数 N_a-N_d 对电导有贡献. 还要注意，当硅中杂质浓度约高于 10^{16}cm^{-3} 时，迁移率随杂质浓度增加而下降。

2.15 采用四探针可测量半导体的方块电阻，而不必制作与半导体的欧姆接触。四个探针排成一条直线(图 P2.15)，电流通过最外边的两个探针。里边两个探针上的电压降用高阻抗电表测量，流过其中的电流可忽略不计。如果探针间距比样品的厚度大而又比它的表面尺寸小，则方块电阻为

$$R_\square=\frac{\pi}{\ln 2}\times\frac{V}{I}=4.53\times\frac{V}{I}$$

(对小的或厚的样品要用到校正因子。[24])

如果四探针电流 $I=1\text{mA}$，在非常高电阻率的 p 型硅片上的扩散浓度为 $N'=10^{12}\text{cm}^{-2}$ 的磷原子时，探针上测到的电压 V 是多少? 假定结深 $x_j=1\mu\text{m}$，μ_n 为与平均磷原子浓度有关的迁移率。

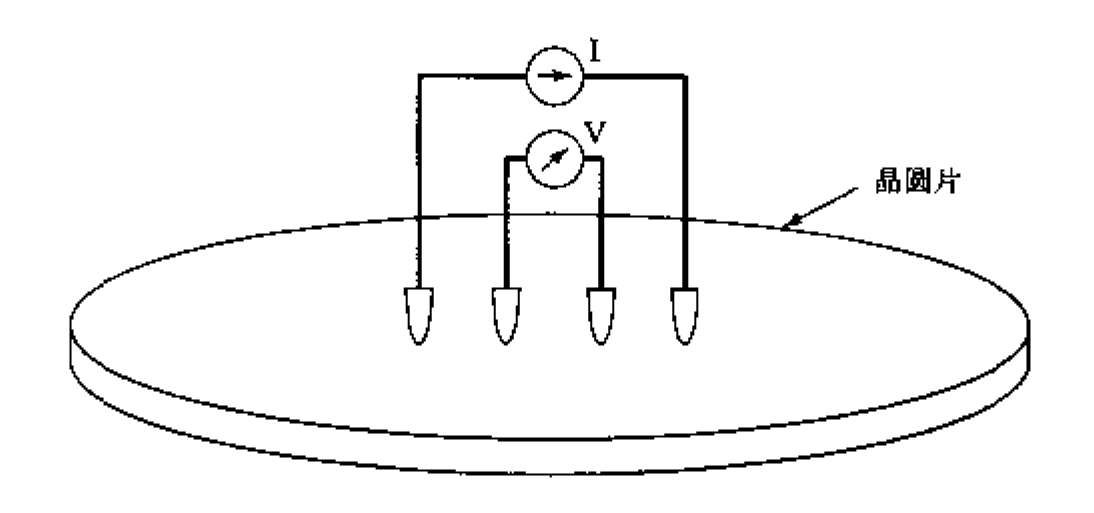


图 P2.15

2.16 已知在电阻率为 5Ω · cm 的 n 型硅衬底中注入剂量为 10^{12}cm^{-2} 的硼离子，注入能量为 100keV（R_p = 290nm，ΔR_p = 70nm），然后在 1000℃扩散 2 小时（$D = 2\times10^{-14}$cm^{2}s^{-1}）。

（a）硼的峰值浓度是多少？刚注入后 p 型区的宽度是多少？

（b）在随后的退火工艺完成后，硼的峰值浓度是多少？

2.17$^+$ 剂量为 N'cm^{-2} 的杂质离子被注入到背景杂质浓度为 C_B 的硅衬底中形成 pn 结。

（a）证明垂直结深为

$$x_j = R_p + \Delta R_p\left[2\ln\left(\frac{N'}{\sqrt{2\pi}\,\Delta R_p C_B}\right)\right]^{1/2}$$

（b）在硼的掺杂浓度 $N_a = 10^{16}$cm^{-3} 的硅衬底中注入剂量为 10^{15}cm^{-2}、能量为 60keV 的砷离子，计算结深 x_j。

2.18* 电阻率为 500μΩ · cm 的某一多晶硅互连线宽度为 5μm、厚度为 0.5μm。电流通过 1mm 长的互连线给电容器充电，电容器的极板夹在 100nm 厚的二氧化硅层的两边，面积为 0.1mm × 0.5mm。计算电阻和电容串联连接的 RC 时间常数？（这个问题中给出的多晶硅的电阻率是可能得到的最小值；这个问题表明了在 VLSI 中采用多晶硅作为互连的局限性。）

2.19 假定在硅中嵌入 6.55μm 厚、具有相反导电类型且浓度均匀的一层硅材料以形成集成电路的电阻（图 P2.19）。

在（i）$N_d = 10^{16}$cm^{-3} 和（ii）$N_a = 10^{16}$cm^{-3} 两种情况下，计算：

（a）25℃时，电极间的电阻分别为 100Ω，1kΩ 和 10kΩ 时，L 和 W 之间的关系。

（b）如果每一电阻消耗的功率为 10mW，最大功耗为 1μW/μm^3，实际尺寸是多少。

（c）25℃左右时的电阻温度系数（TCR）。电阻 R 的 TCR 定义为 $(1/R)(\partial R/\partial T)\times 100$（用每度的百分比表示）。

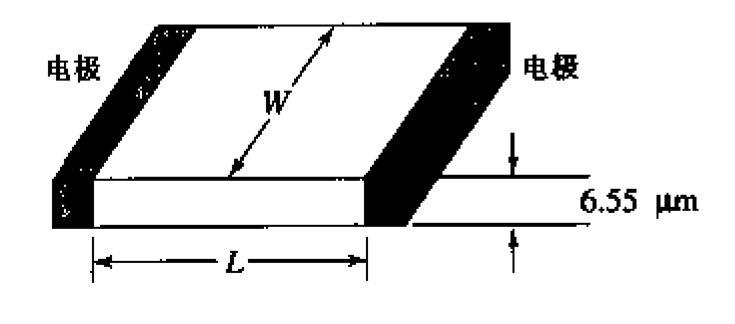


图 P2.19

2.20* 假定在硅片表面得到了如图 P2.20A 所示的阶跃掺杂分布。

（a）不使用平均迁移率来计算出方块电阻。

（b）在经过了若干工艺后，假定杂质沿 x 轴的分布变为图 P2.20B。假定在 $x = 0$ 和 $x = 4$μm 之间以某种

方式能均匀引入其他的杂质原子。应该加入哪种类型的杂质(施主或受主),加入的浓度为多少时才能使均匀掺杂分布(图 P2.20B)时的方块电阻与图 P2.20A 中的方块电阻相等?

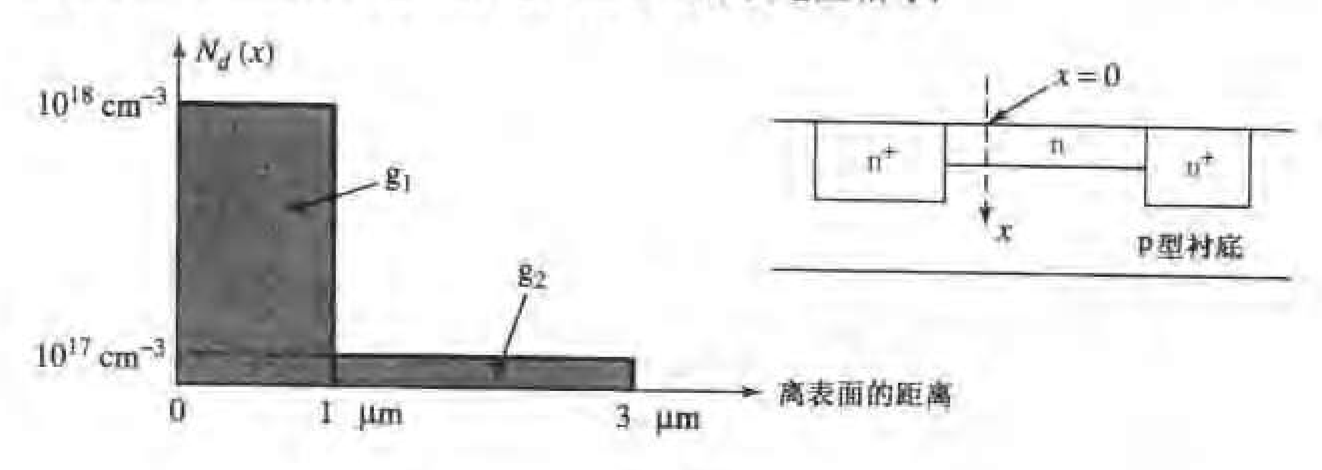


图 P2.20A

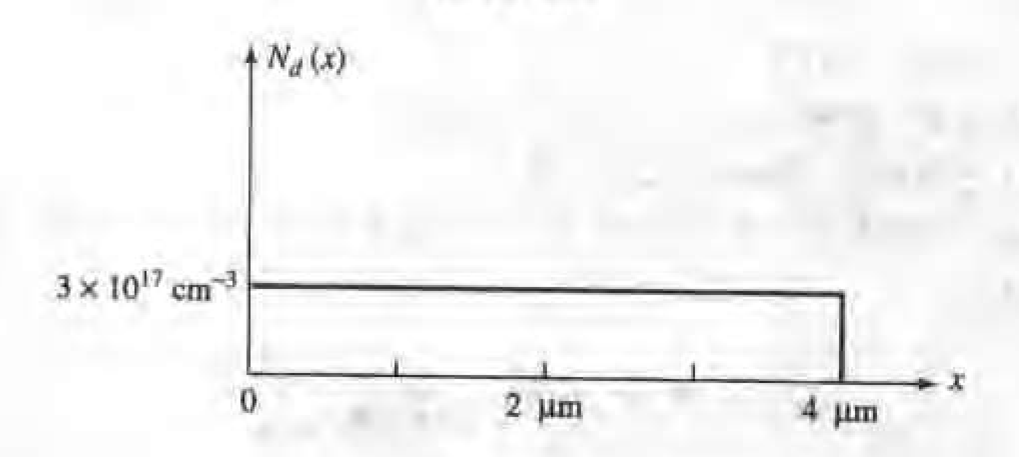


图 P2.20B

2.21* 在轻掺杂的 n 型硅片中进行受主扩散以形成集成电路中的电阻。引入的受主分布为 $N_a = N_s \text{erfc}(x/\lambda)$,其中 $N_s = 10^{18}\text{cm}^{-3}$,$\lambda = 0.05\mu\text{m}$。

(a)证明在电阻表面,任意方块图形的电阻值近似为

$$R_\square \approx \left(q\mu \int_0^\infty N_s \text{erfc}\left(\frac{x}{\lambda}\right)\text{d}x\right)^{-1}$$

请解释在哪一步进行了近似。这个表达式可采用分步积分,$R_\square$值可表示为

$$R_\square = \frac{\sqrt{\pi}}{q\mu N_s \lambda}$$

(b)推导这一公式。

(c)如果电阻宽度为 4μm、长度为 200μm,它的阻值大约是多少?

(d)如果最小线宽是 1μm,最小间距也是 1μm,那么在 20μm × 7μm 的表面积上可以制作的最大电阻值是多少?画出电阻的图形(考虑每一拐角处方块只有 65% 是有效的)。在本题中忽略横向扩散。

第3章　金属-半导体接触

构成集成电路的大部分电子器件是用金属-半导体接触来连接的，而且所有的集成电路与其他电子系统的信号传输也是通过金属-半导体接触进行的。我们将看到这些接触的性质会有相当大的变化，而且为了理解这些变化，必须考虑一些因素。本章我们只讨论金属与硅的接触，我们将首先考虑热平衡下金属和半导体密切接触的普遍特性。了解这一平衡态下的概念非常重要，在对器件的讨论中将多次用到这些概念。因为它们是半导体 pn 结理论的基础，同时也是半导体-绝缘体界面和金属-绝缘体界面理论的基础。

将平衡态下的理论应用于金属-半导体系统，可提供一种简单的关于不同的金属和半导体欧姆接触和整流接触的理论（Schottky 理论）。然而，Schottky 理论在许多情况下是不够的，因为它没有进一步考虑金属-硅系统中固体界面的真实特性，因此还需要讨论表面态的起源及其重要影响。最后将介绍金属-半导体接触的几种应用，重点讨论快速逻辑电路中的 Schottky 箝位二极管。

3.1　电子系统中的平衡

金属-半导体系统

构造金属-半导体接触的能带图对分析本节的问题非常有用。首先金属和半导体代表两种电子能态的系统。利用第1章中的概念知道，这些系统的允许态在能量低于 Fermi 能级时是几乎完全被占据的，而高于 Fermi 能级时几乎是全空的。当金属和半导体彼此很远时，没有相互作用，两个系统的电子状态和它们的 Fermi 能级是独立的。假设金属为系统1，半导体为系统2，每一个系统具有单位能量的态密度 $g(E)$。在这些 $g(E)$ 态中，$n(E)$ 为填充态密度，$v(E)$ 为空态密度（变量都是能量的函数，因此括号中为 E）。图3.1给出了定性描述方法。Fermi-Dirac 分布函数将 $v_{1,2}(E)$、$n_{1,2}(E)$ 和 $g_{1,2}(E)$ 联系起来[1]

$$f_{D1,2} = \frac{1}{1 + \exp[(E - E_{f1,2})/kT]} \tag{3.1.1}$$

填充态密度为

$$n_{1,2} = g_{1,2} \times f_{D1,2} \tag{3.1.2}$$

空态密度为

[1] 关于 Fermi-Dirac 统计的更详细的讨论表明 f_0 与允许电子态的特定性质有关，因此可能会有不同的形式。这种不同的形式通常在等式3.1.1中表现为指数项前有因子2或1/2。本书将不关心这一细节。

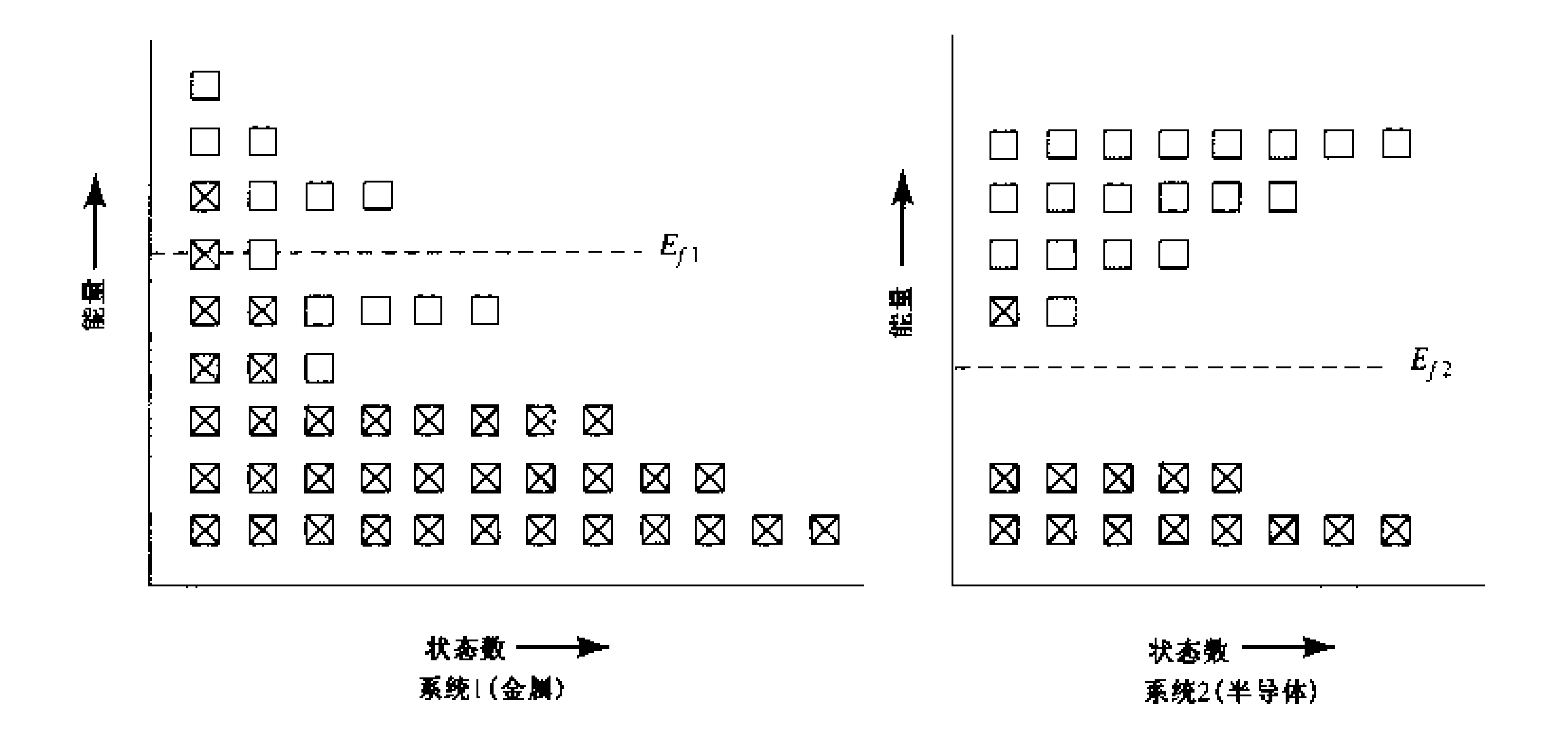


图3.1 两种孤立材料的允许电子能态系统,×表示的状态是填充的;未标注表示空态。系统1定性描述了金属;系统2定性描述了半导体

$$v_{1,2} = g_{1,2} \times (1 - f_{D1,2}) \tag{3.1.3}$$

等式(3.1.1)至(3.1.3)中忽略了表示能量依赖关系的括号。

现在考虑将这两个孤立系统密切接触,两者开始相互作用,电子会在它们之间转移。当没有净的电子转移时系统就达到平衡。正如1.1节提到的,这并不意味着所有的过程都停止了,而是每一过程与它的反过程进行的速率相同。为了使这一点更清楚,考虑某一给定能量 E_x 下电子在空间的转移。这一平衡在数学上的表达式,要考虑给定能量下电子的传输几率正比于电子浓度 $n(E_x)$,以及可接受电子的空态密度 $v(E_x)$。根据Pauli不相容原理,可接受电子的态密度不是 $g(E_x)$,而是空态密度 $v(E_x)$。与这两个态密度相关的传输几率因子与具体状态的量子性质有关,从系统1至系统2的传输与从系统2至系统1的传输几率相等。因此在热平衡下对任何给定的能量

$$n_1 \times v_2 = n_2 \times v_1 \tag{3.1.4}$$

将等式(3.1.2)和(3.1.3)代入等式(3.1.4),有

$$f_{D1}g_1(1 - f_{D2})g_2 = f_{D2}g_2(1 - f_{D1})g_1$$

或

$$f_{D1}g_1g_2 = f_{D2}g_2g_1 \tag{3.1.5}$$

只有在 $f_{D1} = f_{D2}$ 时,或者根据等式(3.1.1),$E_{f1} = E_{f2}$ 时等式(3.1.5)才成立。因此我们确定了任意两系统热平衡条件下的重要性质:它们具有相同的Fermi能级。注意,在上述推导中,对态密度函数 g_1 或 g_2 没有任何限制。无论这些函数的具体特性如何,两系统在热平衡下Fermi能级是相同的。对多于两系统或者单一系统的情况,这一结论更直接的表述为:热平衡状态下,Fermi能级在整个系统中是常数。

3.2 理想的金属-半导体结

3.2.1 能带图

现在开始考虑金属-半导体能带图的特性。就目前的讨论来说，金属和半导体的电子能态最显著的区别在于允许态密度 $g(E)$ 中 Fermi 能级所处的相对位置。在金属中，Fermi 能级位于连续分布的允许态中，而在半导体中，Fermi 能级处的态密度通常可忽略不计。理想的金属和半导体的 $g(E)$ 随能量的变化关系如图 3.2(a) 和 3.2(b) 所示。

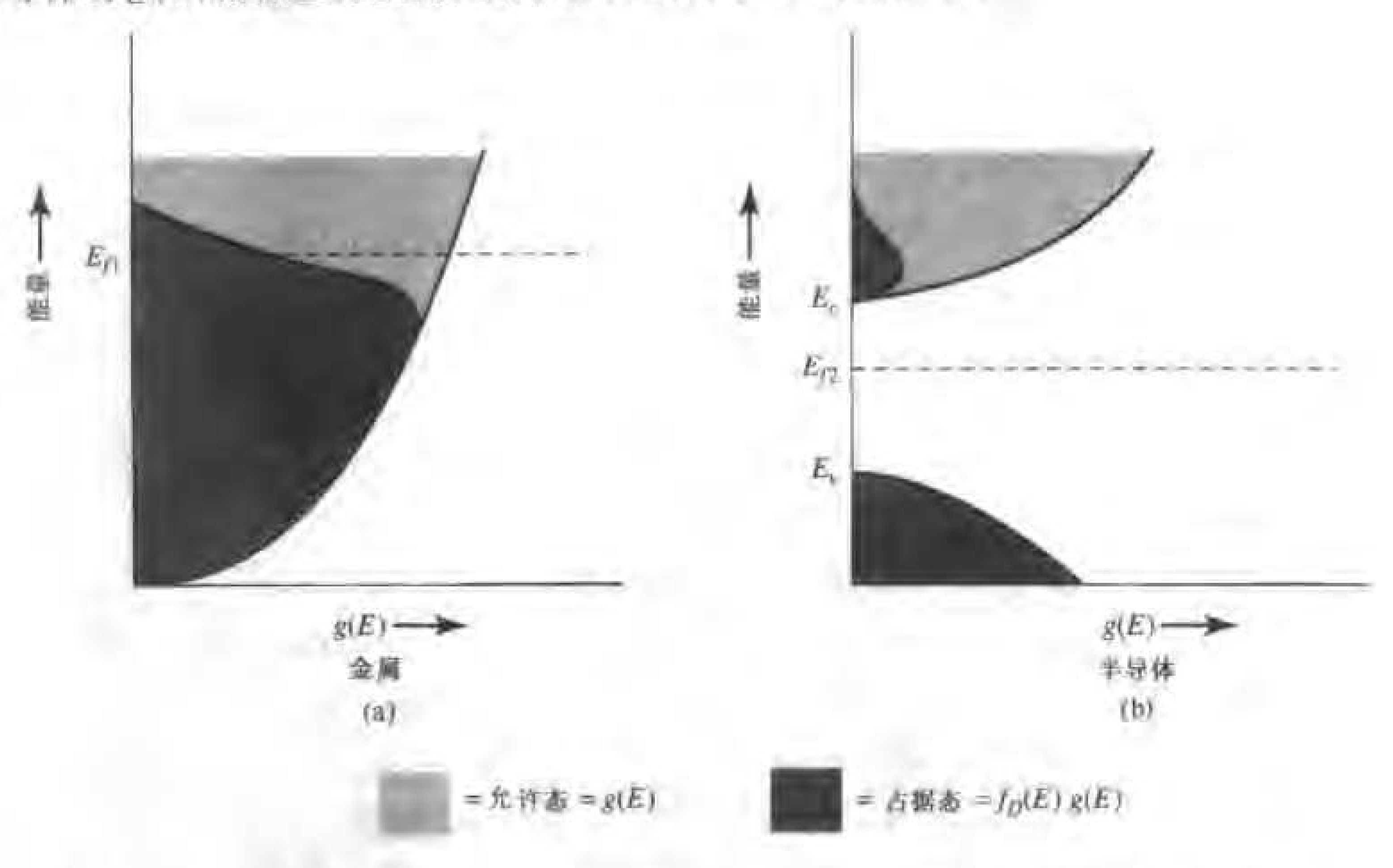


图 3.2 (a)理想金属的允许电子态密度 $g(E)$。注意 Fermi 能级 E_{f1} 位于连续的允许态之中。(b)半导体的允许电子态密度 $g(E)$。Fermi 能级位于导带边和价带边之间。阴影部分的具体含义见图注

图 3.1[②] 中讨论得出的结论对于图 3.3(a) 和 3.3(b) 表示的金和硅这两种特定的材料来说显然是成立的。注意这些图与图 3.2 不同，图 3.3 表示的是体材料中能量随位置变化的能带图，而不是允许态密度随能量变化的能带图。这两类能带图都经常用到，注意不要混淆。

这些图中出现了一些新的变量，首先需要指定能量的参考点，比较方便的是取真空能级或自由电子能量 E_0。E_0 表示电子脱离给定材料所需要的最小能量。E_0 与 E_f 之差称为功函数，通常用符号 $q\Phi$ 表示，单位为能量。对于特定的材料，功函数也经常用 Φ 表示，单位为伏特。半导体的 E_0 与 E_f 之差是半导体中掺杂浓度的函数，因为 E_f 在 E_c 和 E_v 之间的位置是随掺杂浓度变化的。特定材料的真空能级与导带边之差是常数。这个量称为电子亲和能，通常以能量为单位，表示为 $q\chi$。许多材料给出的亲和势是以伏特为单位的 χ 值。

② 译者注：原文误为图 1.3。

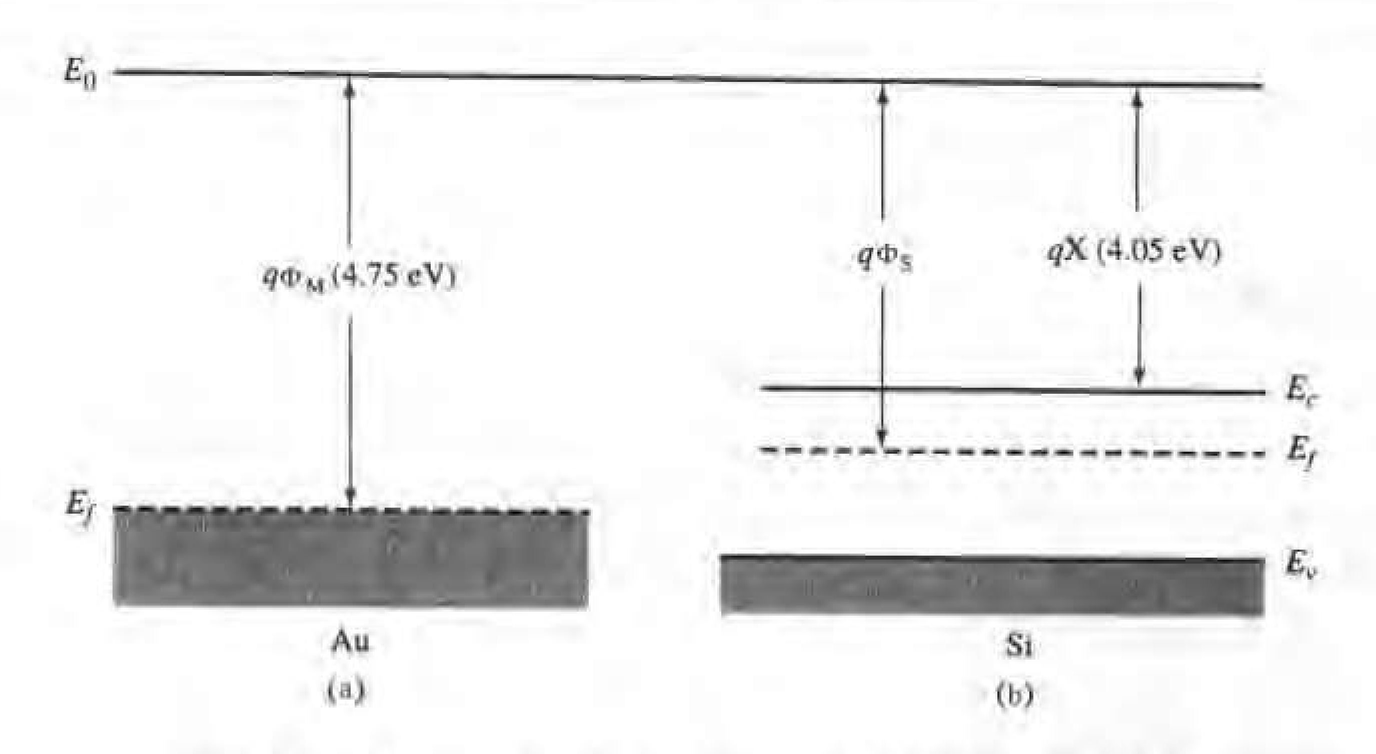


图 3.3 金属材料金和半导体材料硅中的相关能级。图中金属只给出了的功函数 $q\Phi_M$，而半导体给出了功函数 $q\Phi_S$，电子亲和能 qX_S，以及带隙(E_c-E_v)

选择 E_0 为共同的能量参考点，如果 $\Phi_M<\Phi_S$ 且两种材料没有接触，金属中电子的平均总能量将高于半导体中的电子的平均总能量。相反，如果 $\Phi_M>\Phi_S$，半导体中的电子平均总能量将大于金属中电子的平均总能量。为了讨论方便，考虑 $\Phi_M>\Phi_S$ 的情况。当两种材料接触时，平均能量的差异将导致电子从半导体向金属转移。

$\Phi_M>\Phi_S$ 时，另一种确定平衡态的方式是利用 3.1 节的概念。等式(3.1.5)左侧正比于从系统 1 流向系统 2 的电子流，等式的右侧表示相反的电子流。很容易看出如果 $f_{D2}>f_{D1}$ 或者 $E_{f2}>E_{f1}$ 时，净的电子流动是从系统 2 流向系统 1。电荷将持续转移直至建立起热平衡，此时金属和半导体具有相同的 Fermi 能级。在平衡态下，半导体失去电子，相对于金属带正电。

为了构造金属和半导体在平衡条件下的能带图，还需要注意两个问题。首先，真空能级 E_0 必须是连续的。E_0 表示电子刚能获得自由时的能量，因此必须在空间上是连续和单值的。如果不是，我们可以设想一种在平衡态下做功的方式：考虑 E_0 突变的位置，如果在较高的 E_0 处从材料中发射一个电子，然后在无限小距离内的较低的 E_0 处重新吸收这个电子，所做的功不相等，这显然是不合理的。其次，电子亲和能与禁带宽度一样是与晶格结构有关的量，因此对给定的材料是常数。考虑这三个因素：E_f 相等，E_0 连续，半导体中 X 为常数，我们就可以画出金属-半导体系统的能带图的普遍结构。如图 3.4(a)为 n 型半导体的 $\Phi_M>\Phi_S$ 情况下的能带图。

从图 3.4(a)可看出界面处的允许能态有一突变，这一突变的大小为 $q\phi_B$，单位为电子伏特

$$q\phi_B=q(\Phi_M-X) \tag{3.2.1}$$

图 3.4(a)的结边界处的半导体中，位于带边(E_c 和 E_v)的电子能量高于半导体体内的电子能量，这是负电荷从半导体向金属转移的结果。由于电荷的转移，在结附近存在一个电场，带边的电子势能比半导体体内的电子势能高。自由电子在结附近耗尽，E_c 和 E_f 在表面的间距比体内大。

在考虑金属-半导体结的电子特性之前我们注意到，迄今为止的讨论都是基于两种材料的

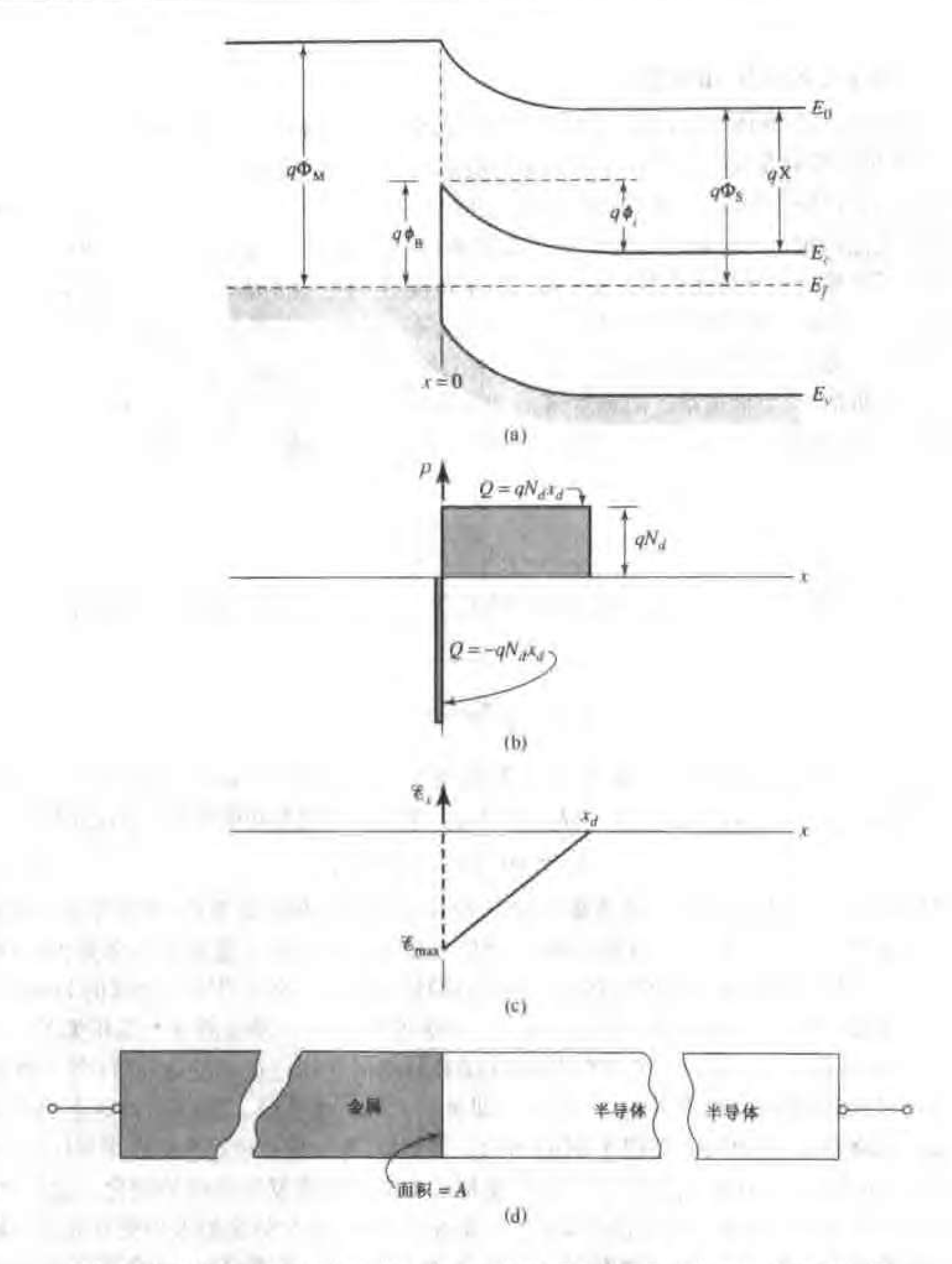


图3.4 (a)金属-半导体整流接触(Schottky势垒)理想的平衡态能带图(能量随位置的变化)。物理界面位于 $x=0$ 处;(b)理想的金属-半导体结的电荷分布。金属表面负电荷近似为一 delta 函数。耗尽近似下正电荷完全是由离化施主(这里假定在空间分布为常数)组成的;(c)理想的金属-半导体结的电场分布

基本能带结构在表面处没有发生变化。用这种理想化的模型将得到一些有用的结果,但是下面有必要更仔细地考虑表面处的条件,以便构造出非理想化的能带图。

3.2.2 电荷、耗尽区和电容

理想的金属-半导体结的电荷与电场分布示于图3.4(b)和3.4(c)。假定金属是完美导体,从半导体转移到金属的电荷位于金属表面的薄层内。理想的n型半导体中,正电荷可由离化的施主或自由空穴组成,电子提供负电荷。在绘制图3.4(b)和3.4(c)时,需要对半导体中的电荷做几点假定:首先,假定各处的自由空穴浓度很小,可以不考虑;其次,从界面到 $x=x_d$ 平面处电子浓度远小于施主浓度;当 $x>x_d$,施主浓度 N_d 等于电子浓度 n。这些假设通常称为耗尽近似。尽管耗尽近似并不严格精确,但足以推导一些非常有用的关系式。在第4章我们将再次考虑平衡态下结附近的空间电荷分布,并且更详细地讨论耗尽近似。

在耗尽近似下,空间电荷区的宽度为 x_d,电场的大小(考虑半导体中掺杂浓度为常数的情况)随位置线性变化(图3.4(c))。电场最大值位于界面处,根据Gauss定理,它的值由下式给出

$$\mathscr{E}_{\max} = \frac{-qN_d x_d}{\epsilon_s} \tag{3.2.2}$$

ϵ_s 为半导体的介电常数。空间电荷区上的电压将等于图3.4(c)中电场曲线包围下的面积的负值

$$\phi_i = -\frac{1}{2}\mathscr{E}_{\max} x_d = \frac{1}{2}\frac{qN_d x_d^2}{\epsilon_s} \tag{3.2.3}$$

后面会经常用到以 ϕ_i 表示的 x_d。从等式(3.2.3)可导出 $x_d=\sqrt{2\phi_i\epsilon_s/qN_d}$。从图3.4(a)可看出内建电势 ϕ_i 等于 $\Phi_M-\Phi_S=\Phi_M-X-(E_c-E_f)/q$。半导体中的空间电荷 Q_s(单位面积)为

$$Q_s = qN_d x_d = \sqrt{2q\epsilon_s N_d \phi_i} \tag{3.2.4}$$

外加偏置 到目前为止,一直考虑的是金属-半导体结的热平衡条件,现在考虑外加偏压的非平衡条件。在图3.4(a)中看到金属-半导体界面允许态的电子能量有一突变,这一能量变化使得电子从金属到半导体的转移比相反方向的转移困难。靠近界面处金属的Fermi能级和半导体导带之间有一势垒 $q\phi_B$(图3.4(a))。一级近似下,这一势垒高度与偏压无关。参考图3.4(c)可以看到靠近金属一边加在以delta函数分布的空间电荷区的电压降(等于电场 $\mathscr{E}$ 曲线与坐标轴之间的面积)在平衡态时为零,即金属上没有电压降。空间电荷区上的全部电压降(ϕ_i)都降落在半导体内,如图3.4(a)所示。类似地,外加偏压也完全降在半导体内,半导体内的电压降将与 ϕ_i 不同,总的能带弯曲程度相对于平衡态将发生相应的变化。这样,半导体内导带边的电子要到达金属一边会受到一个势垒的阻碍,这个势垒的大小受外加偏压影响将偏离其平衡态的值 $q\phi_i$。当金属相对于半导体加正偏压时,势垒降低,加负偏压时势垒增大。这两种情况下的能带图分别示于图3.5(a)和3.5(b)。因为这些图对应于非平衡状态,图中没有统一的Fermi能级。电子从Fermi能量高的区域流向Fermi能量低的区域。当然,电流方向与电子流动的方向相反。

为了研究偏压对势垒的影响,考虑半导体接地,正向偏压对应于金属电极接正电压。外加偏压为 V_a,偏压的极性如图3.5(c)所示。反向偏压对应于金属接负电压($V_a<0$)。如果金属-半导体结加负偏压,降落在空间电荷区的电压将增加到(ϕ_i-V_a),那么半导体中的空间电荷密度将从平衡态的值(等式(3.2.4))增至

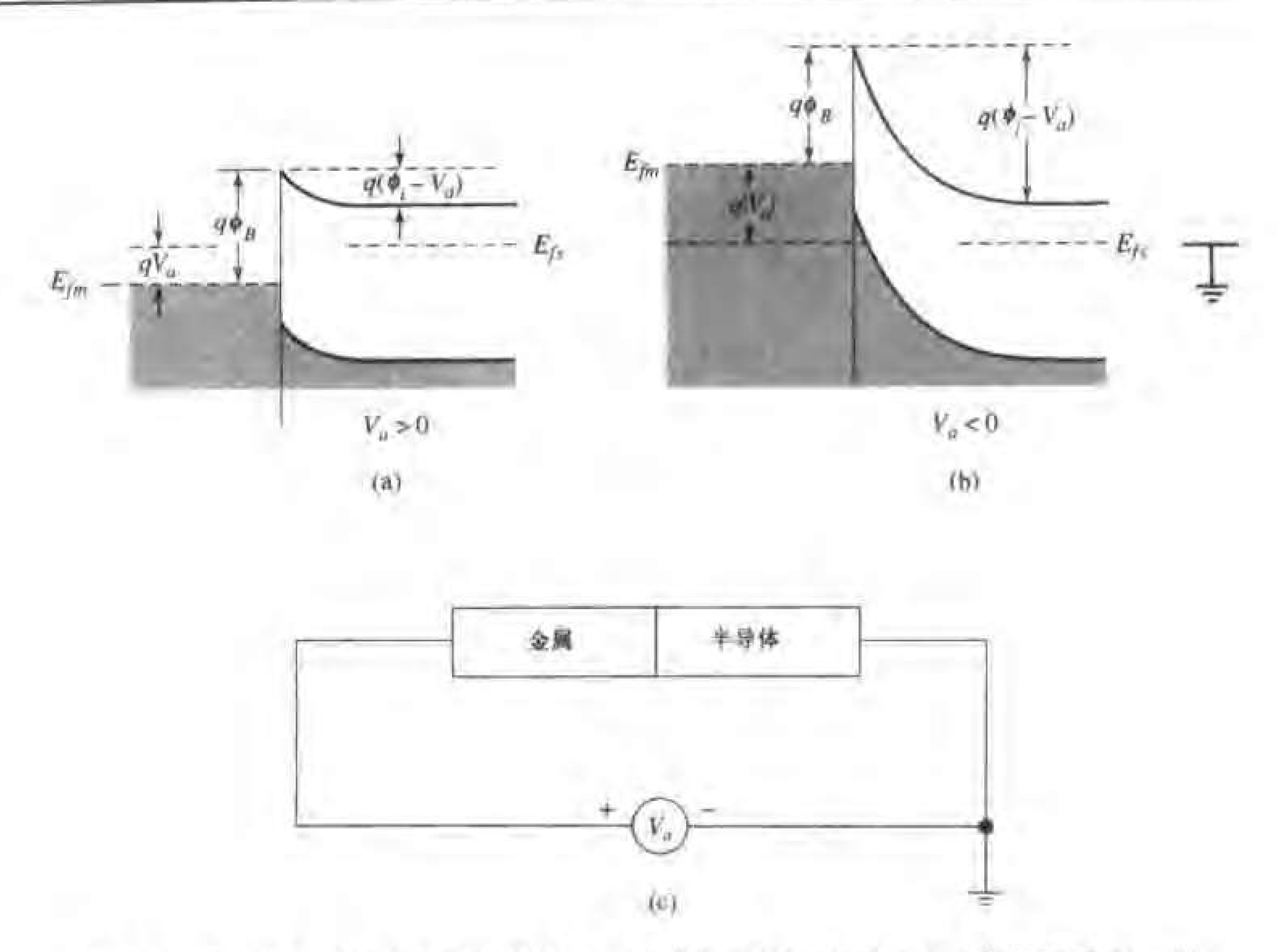


图3.5 理想状态下金属-半导体结在(a)正向偏压下($V_a>0$)和(b)反向偏压下($V_a<0$)的能带图(能量随位置的变化),其中以半导体作为参考点(电压接地),如(c)所示。图中没有给出两种情况下的真空能级

$$Q_s = \sqrt{2q\epsilon_s N_d(\phi_i - V_a)} \tag{3.2.5}$$

将交流小信号叠加到固定的直流偏压 V_a 上,根据等式(3.2.5),结将表现出电容特性

$$C = \left|\frac{\partial Q_s}{\partial V_a}\right| = \sqrt{\frac{q\epsilon_s N_d}{2(\phi_i - V_a)}} = \frac{\epsilon_s}{x_d} \tag{3.2.6}$$

等式(3.2.6)中的 C 表示单位面积的电容,其最终形式是小信号电容 C 的一般表达式。因为 C 表示电荷对电压的微分,它始终可以表示成介电常数除以总的空间电荷区宽度。这一结果将在第4章进一步讨论。

加在结上的总电压降可从等式(3.2.6)中解出来

$$(\phi_i - V_a) = \frac{q\epsilon_s N_d}{2C^2} \tag{3.2.7}$$

等式(3.2.7)的形式表明小信号电容倒数的平方与反向偏压的关系应为直线,如图3.6所示。直线的斜率可用来计算半导体中的掺杂浓度,直线与电压轴的截距等于 ϕ_i。通过测量小信号电容与直流偏压的关系,绘制出与图3.6类似的曲线,这种实验方法通常用于研究半导体的一些特性。实际上,这种方法最大的误差在于对电压轴上的截距 ϕ_i 的计算。根据曲线的斜率能精确算出半导体的掺杂浓度。

当半导体的掺杂浓度随距离变化时,小信号电容的测量也很有用。在这种情况下,空间电荷区分布与图3.4(b)不同,如图3.7所示。对一给定直流(反向)偏压($V_a<0$),空间电荷层

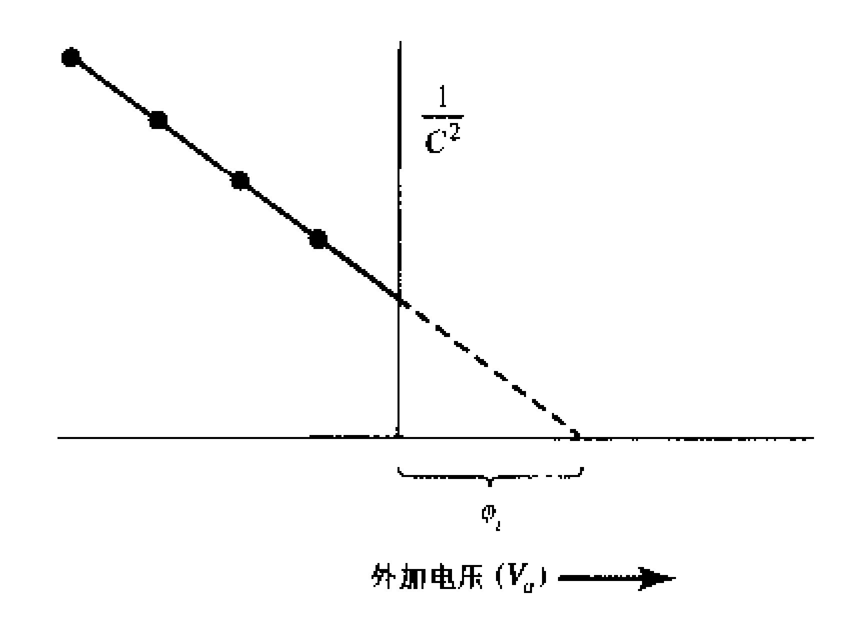


图3.6 理想金属-半导体结的$1/C^2$随偏压的关系曲线

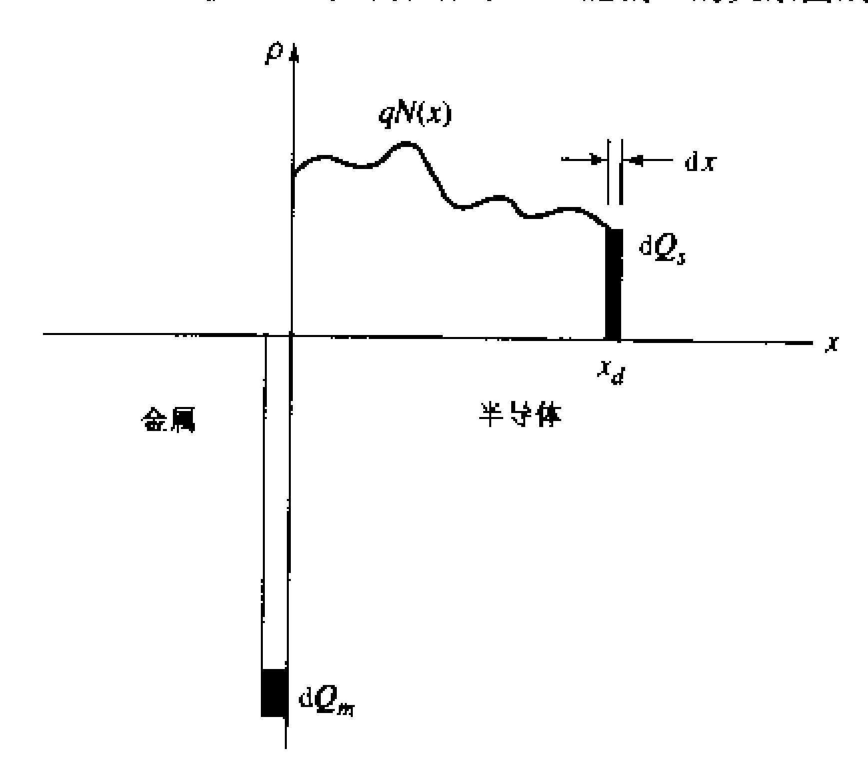


图3.7 半导体非均匀掺杂的金属-半导体结的空间电荷示意图

的宽度为x_d。V_a值的微小增加将引起Q_s的相应的微小的增加,这里

$$Q_s = q\int_0^{x_d} N(x)\,\mathrm{d}x \tag{3.2.8}$$

因此电压的增量$\mathrm{d}V_a$引起Q_s的增量为

$$\mathrm{d}Q_s = qN(x_d)\,\mathrm{d}x = -C\,\mathrm{d}V_a$$

或

$$N(x_d) = -\frac{C}{q\,(\mathrm{d}x/\mathrm{d}V_a)} \tag{3.2.9}$$

x_d为测量小信号电容C时,所加的直流偏压对应的耗尽层宽度。因为小信号电容C可表示成$C=\epsilon_s/x_d$,而微分($\mathrm{d}x/\mathrm{d}V_a$)可写成$\mathrm{d}x/\mathrm{d}V_a=(\mathrm{d}x/\mathrm{d}C)(\mathrm{d}C/\mathrm{d}V_a)=-(\epsilon_s/C^2)\mathrm{d}C/\mathrm{d}V_a$,等式(3.2.9)可重写成

$$N(x_d) = \frac{C^3}{q\epsilon_s(\mathrm{d}C/\mathrm{d}V_a)} \tag{3.2.10}$$

采用$\frac{d(1/C^2)}{dV_a} = -\left(\frac{2}{C^3}\right)\frac{dC}{dV_a}$，等式(3.2.10)可写成更有用的形式

$$N(x_d) = \frac{-2}{q\epsilon_s[d(1/C^2)/dV_a]} \tag{3.2.11}$$

等式(3.2.11)的结果表明$1/C^2$与反向偏压之间的关系曲线的斜率直接反映了空间电荷区边缘的掺杂浓度。这一斜率用$(2/q\epsilon_s)$除将直接给出$N(x_d)$。一些商用半导体分析仪就是利用这一技术来确定掺杂浓度分布的，有些分析仪还可以直接读出等效的掺杂浓度。

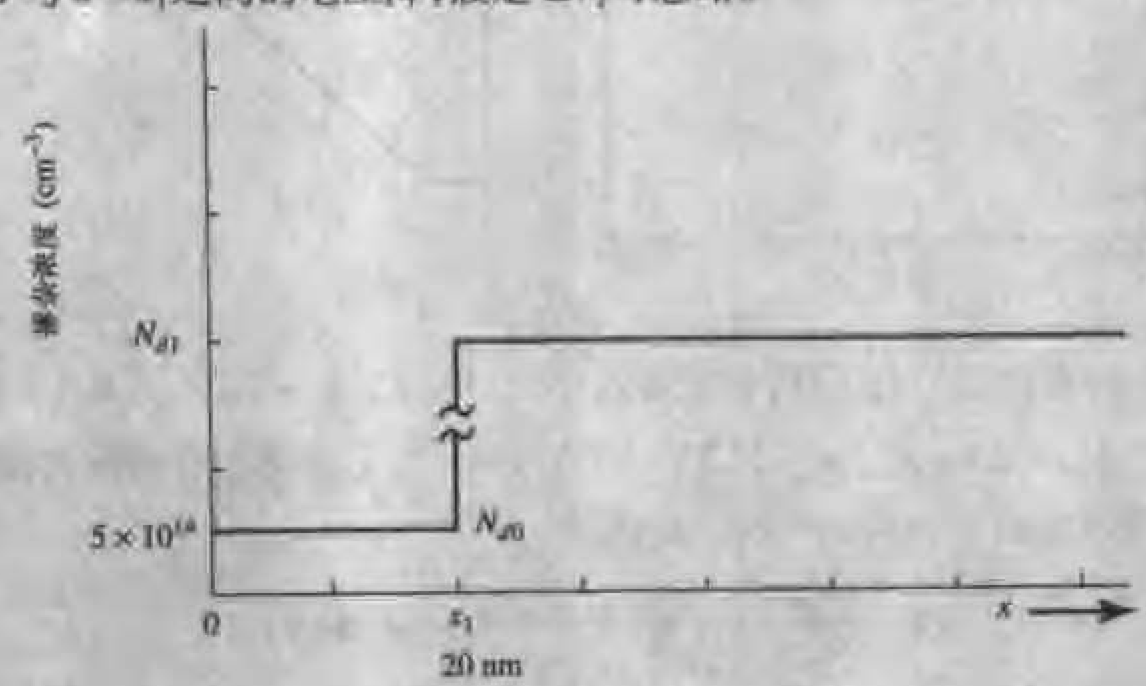
例题　Schottky 势垒二极管

已知金与n型硅的Schottky接触位于$x=0$平面，从$x=0$至$x=x_1$($x_1=20$nm)的硅的表面层掺杂浓度为$N_{d0}=5\times10^{14}\text{cm}^{-3}$。$x>x_1$为高掺杂浓度$N_{d1}(\text{cm}^{-3})$的n型硅，如下图所示。金属-硅接触的内建电势$\phi_i$等于0.50V。假定金的功函数$q\Phi_M$等于4.75eV。

(a)求高掺杂区硅的掺杂浓度N_{d1}；

(b)在热平衡条件下($V_a=0$)画出二极管的电荷密度和电场的分布图；

(c)假定降落在硅表面层的电压ΔV_i可忽略，计算最大的电场值。利用计算出的最大电场值估算$x=0$与$x=x_1$之间的电压降，假定ΔV_i可忽略。



解：由于表面层很薄而且是轻掺杂，可以假定是完全耗尽的，并假定空间电荷区向硅晶体内部延伸。如果这一假定不正确，后面将说明计算误差。从图3.4有

$$q(\Phi_M - \Phi_S) = q\phi_i = 0.50\text{ eV}$$

因此

$$q\Phi_S = 4.75 - 0.50 = 4.25\text{ eV}$$

仍从图3.4可得到

$$q\Phi_S = q\mathrm{X} + (E_c - E_f)$$

或者

$$(E_c - E_f) = 4.25 - 4.05 = 0.20\text{ eV}$$

以及

$$(E_f - E_i) = (E_c - E_i) - (E_c - E_f)$$

这样

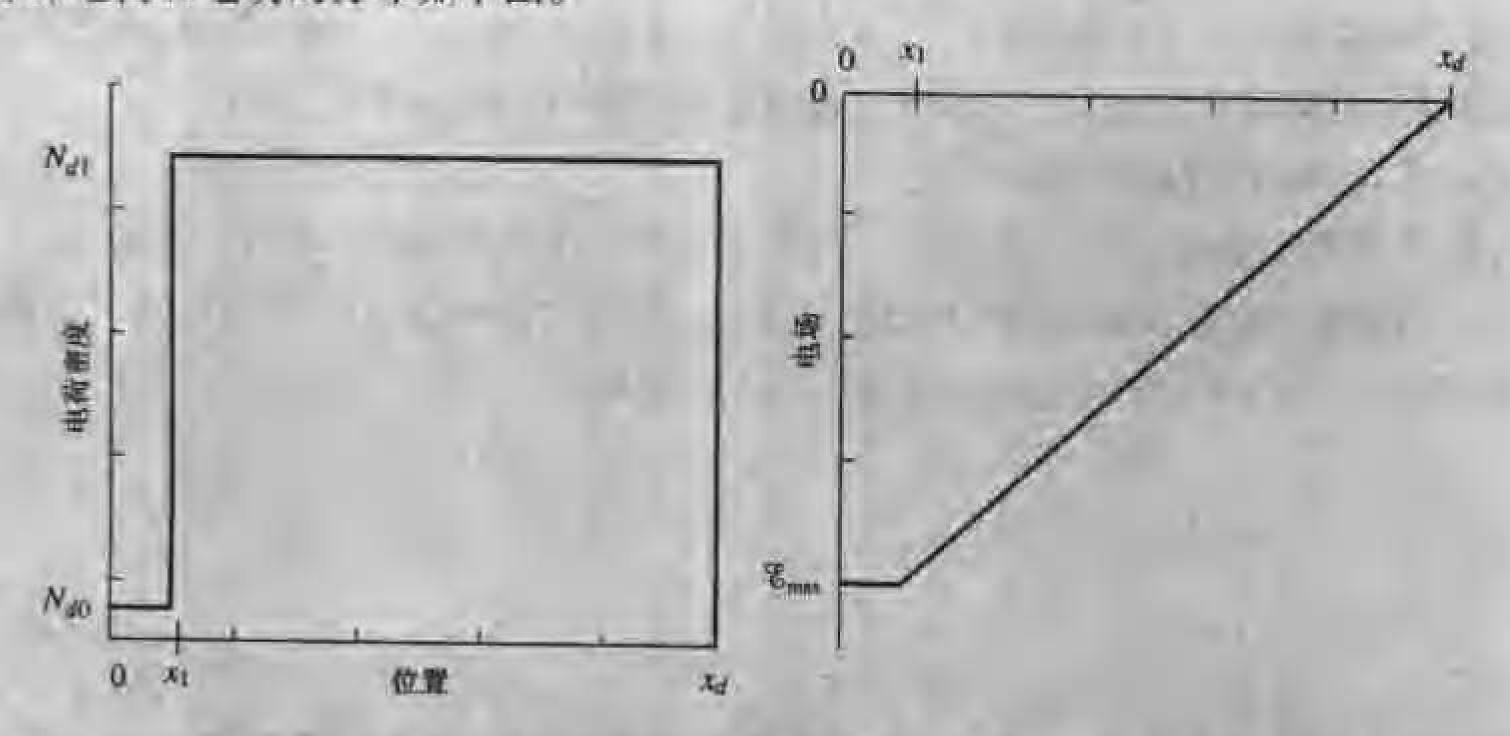
$$(E_f - E_i) = 0.562 - 0.20 = 0.362\ \mathrm{eV}$$

(a)利用等式(1.1.26)可计算出

$$n = N_{d1} = n_i \exp\left(\frac{0.362}{0.0258}\right) = 1.8 \times 10^{16}\ \mathrm{cm}^{-3}$$

计算得出的 N_{d1} 值远大于 N_{d0},证实了表面层完全耗尽的假设。

(b)电荷和电场的分布如下图。



(c)如果忽略表面层的电压降,根据等式(3.2.3),$\mathscr{E}_{max} = -2\phi_i/x_d$,其中 $x_d = \sqrt{2\phi_i\epsilon_s/qN_d}$ $=190\mathrm{nm}$,所以,$\mathscr{E}(x=0) = \mathscr{E}_{max} = -5.27\times10^4\ \mathrm{V/cm}$。为了估算加在表面薄层上的电压降,根据 Gauss 定理可求出这一区域电场的变化

$$\Delta\mathscr{E} = \frac{qN_{d0}}{\epsilon_s}\Delta x = \mathscr{E}(x=0) - \mathscr{E}(x_1) = 154\ \mathrm{V\ cm}^{-1}$$

这样通过表面层的电场可近似看做常数,计算出的 $\Delta V_s = -\mathscr{E}_{max}\times x_1 = 0.105\mathrm{V}$。$\Delta V_s$ 大约高达 ϕ_i 的 20%,应当在不忽略表面层上电压降的情况下重新做计算。

Schottky 势垒的降低[†] 现在重新考虑一级近似的结论,即电子从金属流向半导体时遇到的势垒不随外加偏压变化。反偏时可观察到势垒高度与偏压之间较弱的依赖关系。这一依赖关系与多年前 Walter Schottky 在研究电子向真空发射时所观察到的效应相符合。

下面推导中假定,半导体中靠近界面处电子能量用自由电子近似理论,而金属被视为一导电薄层。在这个模型中,半导体的特征表现在两方面:(1)指定电子的有效质量是 m_n^*,(2)采用半导体的相对介电常数(对硅,$\epsilon_r = 11.7$),而不是真空相对介电常数 1。图 3.8 中给出了平衡条件下和外加偏压条件下的能带图,电场的方向趋向于使电子离开金属表面。图中代表电子能量的函数 $E_1(x)$ 用经典方法推导出;导电金属薄层对电子有镜像力的作用,即相当于具有相反符号的镜像电荷被放置在 $x=0$ 平面左边。当外加电场 $-\mathscr{E}$ 趋向于使电子远离金属表面

进入半导体内时，电子能量 $E_2(x)$ 为

$$E_2 = \frac{-q^2}{16\pi\epsilon_s x} - q\mathscr{E}x \tag{3.2.12}$$

从金属-真空系统的研究中可知，等式(3.2.12)对距离大于几个纳米的情况是准确的。E_2取极大值的平面很容易求得，对应于图 3.8 中的能量 $q\Delta\phi$ 为

$$q\Delta\phi = \sqrt{\frac{q^3\mathscr{E}}{4\pi\epsilon_s}} \tag{3.2.13}$$

根据这一模型，电子流出金属所要越过的势垒高度 $q\phi_B$随 $q\Delta\phi$ 增加而降低。电流与这一势垒高度呈指数关系，因为金属中按 Boltzmann 分布的电子，只有在能量高于势垒极大值时才能越过势垒。因此反偏下从金属流出的电流为

$$J = J_0 \exp\frac{\sqrt{q^3\mathscr{E}/4\pi\epsilon_s}}{kT} \tag{3.2.14}$$

利用耗尽近似，将电场 $\mathscr{E}$ 与偏压 V_a 和内建电势 ϕ_i联系起来

$$\mathscr{E} = \sqrt{\frac{2qN_d}{\epsilon_s}(\phi_i - V_a)} \tag{3.2.15}$$

等式(3.2.14)和(3.2.15)表明，因为 Schottky 势垒的降低，高反向偏压下电流与电压的 4 次方根呈指数关系。尽管实际中有时会观测到这种指数依赖关系，但是空间电荷区的自由载流子的反向产生电流可能大于这里考虑的分量。在此情形下，反向电流与偏压的依赖关系将减缓。在第 5 章讨论 pn 结的电流时将发现，载流子的产生流是反向电流的一个来源。

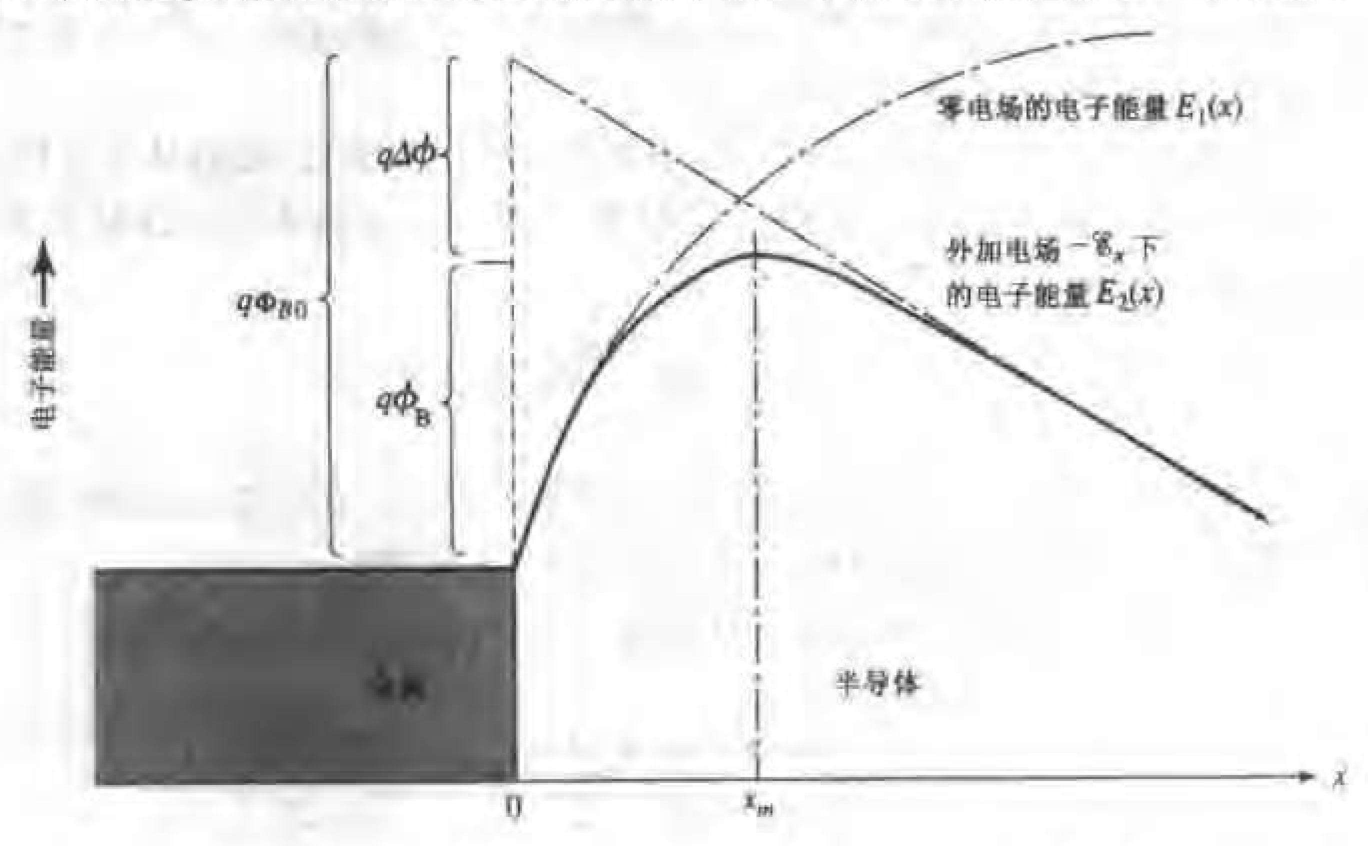


图 3.8　热平衡状态[$E_1(x)$]和外加偏压 $-\mathscr{E}_x$下[$E_2(x)$]靠近金属表面的自由电子经典能量图

3.3 电流-电压特性

我们可以定性推出 Schottky 势垒二极管的电流电压基本关系。这将有利于我们了解金属-半导体系统的平衡态特性,因此在更精确计算 I-V 特性前先进行定性推导。

我们的推导从图 3.4 中热平衡态下的能带图开始。在平衡态,电子从金属越过势垒进入半导体的几率与电子从半导体越过势垒进入金属的几率相等。从第 1 章关于扩散的讨论可知,晶体中的电子因具有热能而不断地运动。例如习题 1.13 证明了处于热运动的自由载流子浓度 n_0 在任意方向上引起的电流密度等于 $-qn_0v_{th}/4$。当然热平衡时这一电流密度与相反方向的电流密度相互抵消,因此净电流为零。将这一概念用于图 3.4 中能带图的边界平面,我们将看到有电子从半导体进入金属,也有电子从金属进入半导体,两者相互平衡。这些电流与边界处的电子浓度成正比。在半导体中,这个浓度 n_s(根据等式(1.1.21))是

$$n_s = N_c \exp\left(-\frac{q\phi_B}{kT}\right) \tag{3.3.1}$$

将体密度 $n = N_d$ 代入等式(1.1.21),从图 3.4 中可知,半导体内 $q\phi_B = q\phi_i + E_c - E_f$,于是有

$$n_s = N_d \exp\left(-\frac{q\phi_i}{kT}\right) \tag{3.3.2}$$

这样,结处于平衡态时有

$$|J_{MS}| = |J_{SM}| = KN_d \exp\left(-\frac{q\phi_i}{kT}\right) \tag{3.3.3}$$

其中 J_{MS} 和 J_{SM} 为热运动产生的电流密度(单位面积电流),分别表示从金属流向半导体的和半导体流向金属的电流密度,式中 K 是比例常数。

当偏压 V_a 加在图 3.5 所示的结上,半导体中的电位降发生变化,导致从半导体流向金属的电子电流发生变化。如果假定在外加偏压时半导体的表面和体内仍保持热平衡,等式(3.3.2)所表示的 n_s 变为

$$n_s = N_d \exp\left[-\frac{q(\phi_i - V_a)}{kT}\right] \tag{3.3.4}$$

因此,流出半导体的电子电流表达式也以相同的方式进行修正。因为势垒($q\phi_B$)仍保持热平衡时的值,所以从金属到半导体的电子电流不受外加偏压的影响。

将这两个分量相减,可得到外加偏压下从金属流向半导体的净电流表达式

$$\begin{aligned} J &= J_{MS} - J_{SM} \\ &= KN_d \exp\left[-\frac{q(\phi_i - V_a)}{kT}\right] - KN_d \exp\left(\frac{-q\phi_i}{kT}\right) \end{aligned} \tag{3.3.5}$$

还可写成

$$J = J_0[\exp(qV_a/kT) - 1] \tag{3.3.6}$$

其中定义常数 $J_0 = KN_d\exp(-q\phi_i/kT)$。

等式(3.3.6)通常称为理想二极管方程。正如推导过程所表明的,理想的二极管方程适

用于这种情况，即在外加偏压下，电子越过的势垒在金属一侧不变，而只在半导体一侧改变，从而影响了载流子的热发射电流。尽管更详细的分析将对电流方程有小的修正，等式(3.3.6)中金属-半导体 Schottky 结的电流与电压的基本依赖关系依然成立。理想二极管方程表明当 V_a 为负值时会出现一个饱和电流 $-J_0$，而正向偏压时电流会急剧增加(习题3.10)。

在第5章讨论流过半导体 pn 结的电流时会再次遇到理想二极管方程(等式(3.3.6))。在推导 pn 结电流方程时，将用到第4章中讨论的半导体的能带、电压和掺杂浓度之间的关系。现在读者可以直接跳到第4章，完成第4章和第5章的阅读，再回到第3章学习剩下的关于金属-半导体结的理论。这样无论逻辑关系还是内容的连续性都不会受到影响，这仅仅提供了适应特定读者兴趣的另一种方法。

更详细的关于金属-半导体 Schottky 结的 J-V 特性的讨论表明，饱和电流 J_0 并不完全与外加偏压无关。本节下面的内容将对此进行分析。

3.3.1 Schottky 势垒†

通过对结附近耗尽区的载流子扩散和漂移方程进行积分，可以求出金属-半导体结的电流电压的关系。这种方法首先由 Schottky[1] 提出，其中假定空间电荷区的尺寸足够大，扩散常数和迁移率都有意义，这样耗尽区的宽度至少为几个电子平均自由程，电场强度小于漂移速度饱和时对应的电场。另一种物理方法是由 Bethe[2(a),(b)] 首先提出的，它基于金属的热离子发射理论，即使上面的一些约束条件不成立，也能推导出同样的 J-V 关系[3]。

在一维情况下考虑图3.9中外加偏压下的金属-半导体结中越过势垒的电子流，那么(正如等式(1.2.21))

$$J_x = q\left[n\mu_n \mathscr{E}_x + D_n \frac{\mathrm{d}n}{\mathrm{d}x}\right] \tag{3.3.7}$$

如果电势用 ϕ 表示，利用 $\mathscr{E}_x = -\mathrm{d}\phi/\mathrm{d}x$ 以及 Einstein 关系(等式(1.2.20))可写出

$$J_x = qD_n\left[\frac{-qn}{kT}\frac{\mathrm{d}\phi}{\mathrm{d}x} + \frac{\mathrm{d}n}{\mathrm{d}x}\right] \tag{3.3.8}$$

图3.9(b)为金属接地时 ϕ 与位置的关系。

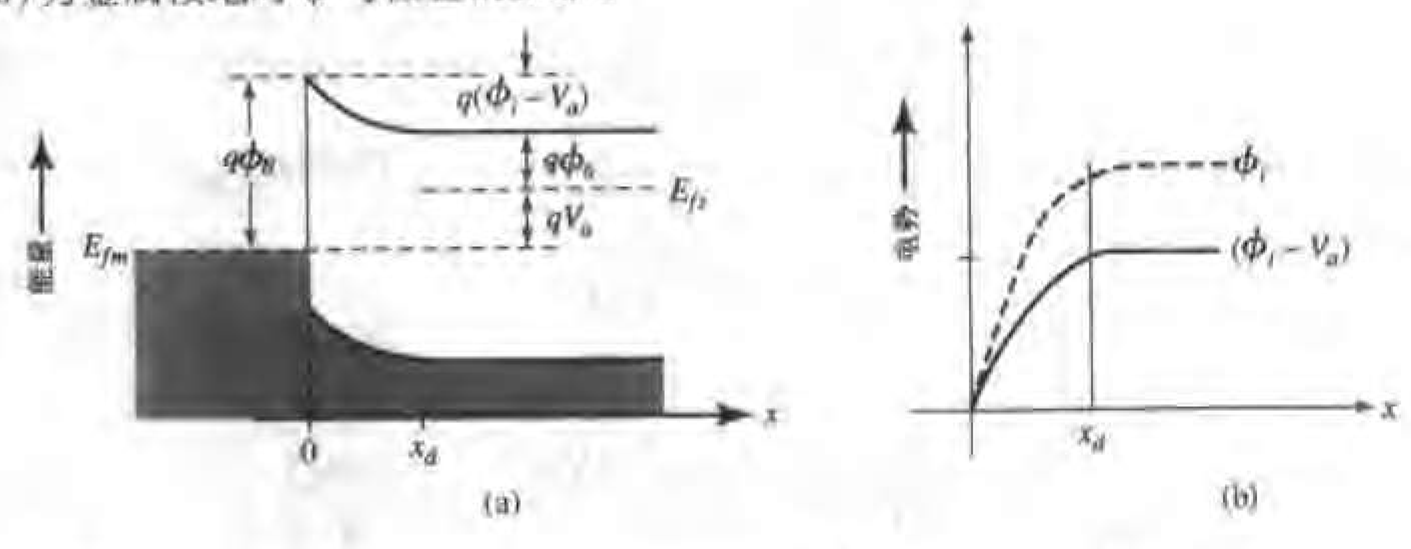


图3.9 (a)正向偏压下金属-半导体整流结的能带图。外加偏压使 Fermi 能级移动：$qV_a = E_{fs} - E_{fm}$。(b)表面耗尽层的电压降减小至 $\phi_i - V_a$

接下来，将等式(3.3.8)写成在耗尽区两端之间的积分形式(图3.9中的 $x=0$ 和 $x=x_d$)。

对等式(3.3.8)两端同时乘以积分因子$\exp(-q\phi/kT)$,则等式右边可直接求出积分。利用耗尽区两端的边界条件,有

$$J_x\int_0^{x_d}\exp\left(\frac{-q\phi}{kT}\right)\mathrm{d}x = qD_n\left[n\exp\left(\frac{-q\phi}{kT}\right)\right]_0^{x_d} \tag{3.3.9}$$

在写等式(3.3.9)时,假定粒子流J_x与位置无关,可以放在积分号的外面。这一假设适用范围很宽。因为电势以金属为参考点,等式(3.3.9)中电压的边界条件为

$$\phi(0)=0 \qquad 和 \qquad \phi(x_d)=(\phi_i-V_a) \tag{3.3.10}$$

正如从图3.9(a)中看到的,$\phi(x_d)$也可写为

$$\phi(x_d)=(\phi_B-\phi_n-V_a) \tag{3.3.11}$$

式中$q\phi_n$等于半导体内的(E_c-E_f)。等式(3.3.9)中还需要确定电子浓度n的边界条件。利用等式(1.1.21),可将边界条件表示成

$$n(0)=N_c\exp\left(\frac{-q\phi_B}{kT}\right)$$

和

$$n(x_d)=N_d=N_c\exp\left(\frac{-q\phi_n}{kT}\right) \tag{3.3.12}$$

将边界条件代入等式(3.3.9)中有

$$J_x=qD_nN_c\exp\left(\frac{-q\phi_B}{kT}\right)\left[\exp\left(\frac{qV_a}{kT}\right)-1\right]\Big/\int_0^{x_d}\exp\left(\frac{-q\phi(x)}{kT}\right)\mathrm{d}x \tag{3.3.13}$$

为了得到电流与电压的关系,ϕ与x的依赖关系必须代入等式(3.3.13)的分母的积分中并求出积分。ϕ与x的这个函数关系决定于靠近结处的半导体的掺杂分布。

前面一直讨论的是电子越过势垒从金属流向均匀掺杂半导体,这一势垒称为Schottky势垒,以纪念第一个对它进行研究的科学家Schottky[1]。对Schottky势垒,在3.2节中采用耗尽近似,得到了耗尽区电势的表达式

$$\phi(x)=\frac{qN_d}{\epsilon_s}x\left(x_d-\frac{x}{2}\right) \qquad (0<x<x_d) \tag{3.3.14}$$

将等式(3.3.14)代入到等式(3.3.13)进行积分,可得到J_x关于V_a的显式解

$$J_x=J_S\left[\exp\left(\frac{qV_a}{kT}\right)-1\right] \tag{3.3.15}$$

其中

$$J_S=\frac{q^2D_nN_c}{kT}\left[\frac{2q(\phi_i-V_a)N_d}{\epsilon_s}\right]^{1/2}\exp\left(\frac{-q\phi_B}{kT}\right) \tag{3.3.16}$$

等式(3.3.16)表明J_S并非与电压无关,因此,等式(3.3.15)中电流与电压依赖的关系是部分隐含在J_S中的。J_S与电压的平方根关系与等式(3.3.15)中的指数关系相比要弱一些,因此可以将等式(3.3.15)的电流与电压关系近似为

$$J_x = J_S'\left[\exp\left(\frac{qV_a}{nkT}\right) - 1\right] \tag{3.3.17}$$

式中，J_S'与电压无关，n 是一个参数，实验上确定的值在 1.02 和 1.15 之间（见习题3.9）。实测的正向偏置的铝-硅 Schottky 势垒的实验结果如图 3.10 所示。测量值与利用等式(3.3.17)从图中拟合出来的 $n = 1.07$。

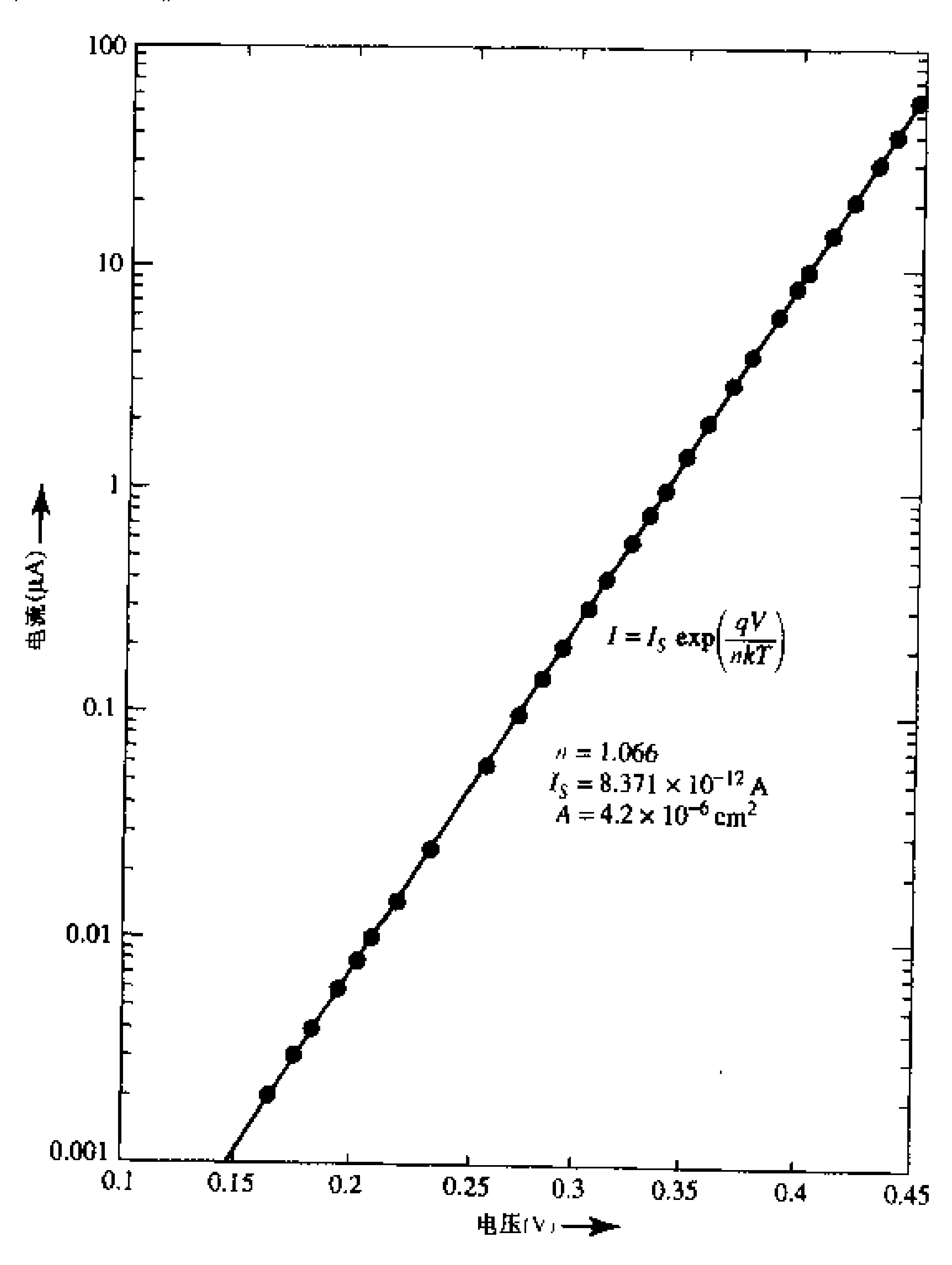


图 3.10　测量得到的铝-硅 Schottky 势垒的电流（对数坐标）电压关系。$I_S = J_S'A$，n 值可通过等式(3.3.17)进行数据拟合得到

前面的分析利用了许多假定，有些已明确指出。在继续讨论前需要指明一个重要的隐含假定，因为该假定在一级近似的器件分析中经常用到。这个假定就是即使有电流流动，系统仍处于准平衡态（即几乎处于热平衡）。在上面的分析中多次用到了这一隐含的准平衡假定，例如，在写等式(3.3.4)时以及调用 Einstein 关系推导等式(3.3.8)时。从逻辑关系上讲，准平衡假定在电流较小的低偏压下更合理，事实也是这样。通常准平衡分析是足够的，因为对这一理论只要做微小的修正就可以扩展到整个感兴趣的电流范围内。当然，对任何假定最终的检验是实验值与预测值是否较好地吻合，如图 3.10 所示。

3.3.2 Mott 势垒[†]

迄今为止只考虑了金属-半导体结中半导体一侧的空间电荷区均匀掺杂的情况，也就是说只考虑了 Schottky 势垒。为了推导其他掺杂分布的电流电压特性，仍可以利用等式(3.3.13)~(3.3.17)，但是耗尽区电位的表达式(等式(3.3.14))必须修正。N. F. Mott 提出了一种有用的金属-半导体结的掺杂情况，即 Mott 势垒，并将它与氧化物二极管的研究联系起来[4]。

Mott 势垒近似的特点是，半导体在靠近金属界面处掺杂较低，在离界面很短的距离内突变为高掺杂。金属界面到高掺杂区的距离很短，在这种意义上实际没有电力线中止在轻掺杂区(这个条件也就是说该距离远小于 Debye 长度 L_D，L_D将在 3.4 节中讨论)。对应的电子能带图如图 3.11 所示。因为轻掺杂区的厚度很薄，该区域的电场可假定为常数。我们还假设没有电力线穿透高掺杂区，因为那里的施主浓度很高。Mott 势垒描述的是这样一种情况，即在结的制备过程中靠近金属电极的区域掺杂发生了变化。也可描述在晶体重掺杂区上方薄的轻掺杂外延层上制备金属电极的情况，后者在设计双极集成电路时会遇到。

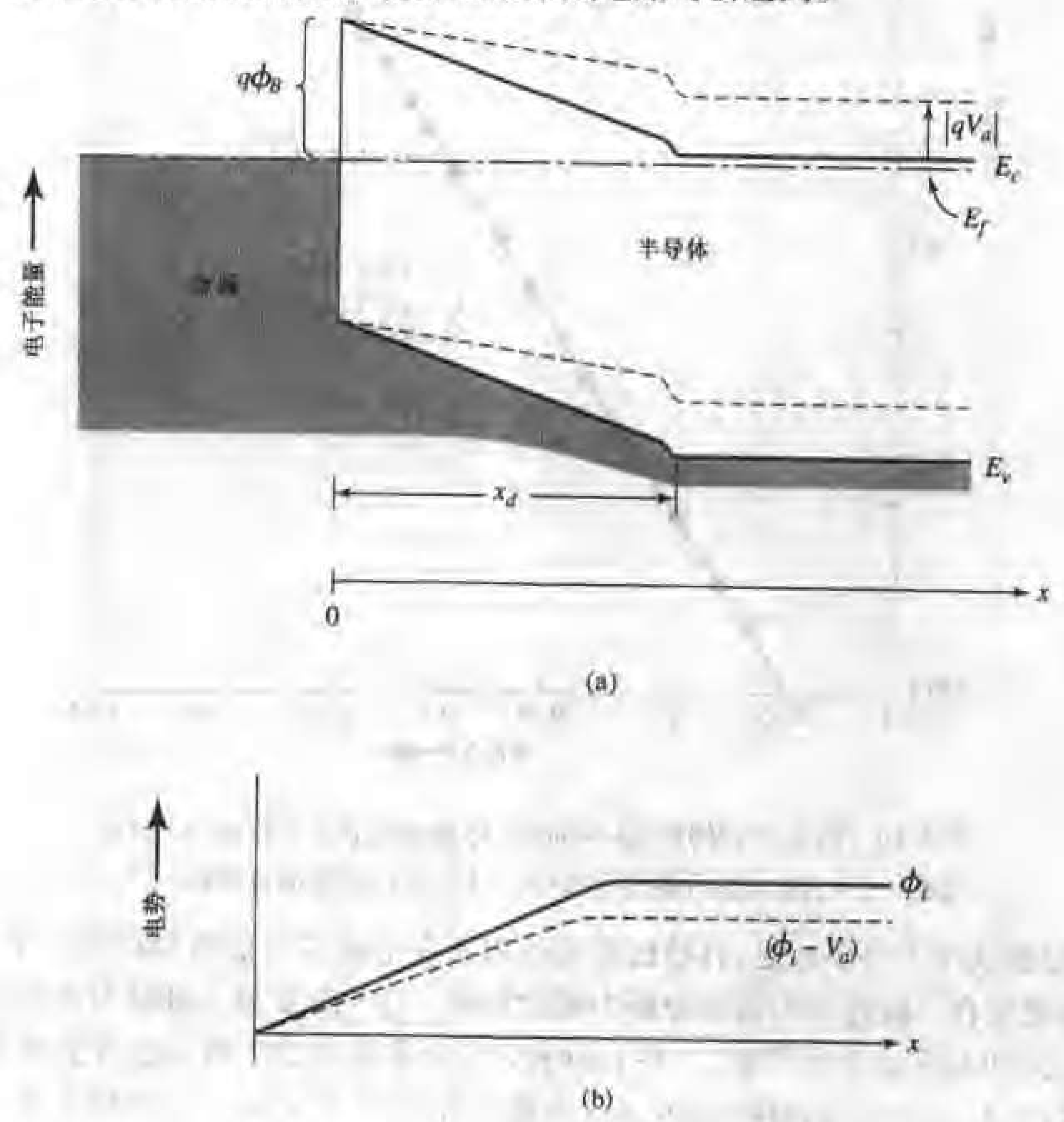


图 3.11 (a) Mott 势垒的电子能量图(表面近似绝缘的区域在 $x=x_d$ 处突变到高电导区)。实线表示热平衡状态；虚线表示加正偏压 V_a 的状态，金属接地。(b) Mott 势垒的电位分布图

为了得到 Mott 势垒的电流与电压关系，与 Schottky 势垒一样，首先写出电势 ϕ 与 x 的关

系，然后对等式(3.3.13)积分。因为轻掺杂区的电场为常数，ϕ 与 x 成线性关系

$$\phi(x) = (\phi_i - V_a)\frac{x}{x_d} \qquad (0 < x < x_d) \tag{3.3.18}$$

将等式(3.3.18)代入等式(3.3.13)，得到与等式(3.3.15)相似形式的结果

$$J_x = J_M\left[\exp\left(\frac{qV_a}{kT}\right) - 1\right] \tag{3.3.19}$$

其中 J_M 比 Schottky 势垒理论推导出的参数 J_S(等式(3.3.16))更强烈地依赖于 V_a

$$J_M = \frac{q^2 D_n N_c(\phi_i - V_a)\exp\left(\frac{-q\phi_B}{kT}\right)}{x_d kT\left\{1 - \exp\left[\frac{-q(\phi_i - V_a)}{kT}\right]\right\}} \tag{3.3.20}$$

Mott 势垒和 Schottky 势垒代表了理想的金属-半导体整流结。在许多情况下，这些理想化的结果就足够了。但是当理想 J-V 关系不够精确时，就需要推导势垒区精确的电位分布并代入等式(3.3.14)或(3.3.18)。总之，正如等式(3.3.15)和(3.3.19)，主要的电压依赖关系包含在指数项里，与测量数据的拟合可采用等式(3.3.17)类似的形式，n 值接近于1。

当 V_a 接近于 ϕ_i 时，推导饱和电流的两个等式(等式(3.3.16)中的 J_S 和等式(3.3.20)中的 J_M)在物理上是不合理的。实际上，如果 V_a 等于 ϕ_i，由于不存在势垒，结中将有很大的电流流过。然而，当实际的二极管加大的正向偏压时，其中相当一部分电压降落在与结串联的半导体体电阻上，因此实际降落在 Schottky 势垒上的正向电压永远不会超过内建电势。串联电阻的影响将在第5章讨论 pn 结电流时详细阐述。

3.4 非整流(欧姆)接触

在讨论金属-半导体结时，考虑靠近金属的半导体区域比体内的多子浓度低，因而从金属流向半导体的电子需越过一个势垒。在这种情况下，偏压主要降落在结区，电流的大小取决于结的性质。相反的情况是，结本身的电阻与半导体体电阻相比几乎可以忽略，这种情况对应的是欧姆接触。尽管欧姆接触的这个定义听起来不很清晰，但强调了一个基本特征：当电压加在具有欧姆接触的器件上时，欧姆接触上的电压降与器件其他部分的电压降相比可忽略。这样，欧姆接触上没有功耗，即使有电流流过，也可认为欧姆接触处于热平衡态。这一特性的一个重要且有用的结果是有电流流过时，欧姆接触处的自由载流子浓度不发生变化，保持热平衡态时的值。

3.4.1 隧道接触

如果势垒对载流子的阻碍作用可忽略，前面介绍的金属-半导体接触就变为欧姆接触。实际上欧姆接触是通过对半导体重掺杂来实现的，这时势垒的宽度 x_d 非常小，从等式(3.2.4)解出 x_d 为

$$x_d = \sqrt{\frac{2\epsilon_s \phi_i}{qN_d}} \qquad (3.4.1)$$

空间电荷区宽度随 N_d 的增加而减小。当势垒宽度接近几个纳米时，发生一种新的输运现象——通过势垒的隧穿。

图3.12(a)表示的是通过非常薄的 Schottky 势垒的隧穿过程。当势垒宽度达到纳米量级，并且金属相对于半导体加负偏压时，即使金属中的电子能量低于势垒高度($q\phi_B$)也可通过隧穿进入半导体导带。同样，当半导体相对于金属加负偏压时，半导体中的电子也可隧穿进入金属(图3.12(b))。当有外加偏置时，由于很多电子参与了这一隧穿过程，所以电流增加很快，因此能发生隧穿的金属-半导体接触具有非常小的电阻，实际应用中的欧姆接触就是这样制备的。为了得到非常薄的势垒，必须对半导体进行重掺杂，直至达到简并(即 Fermi 能级进入价带或导带)。

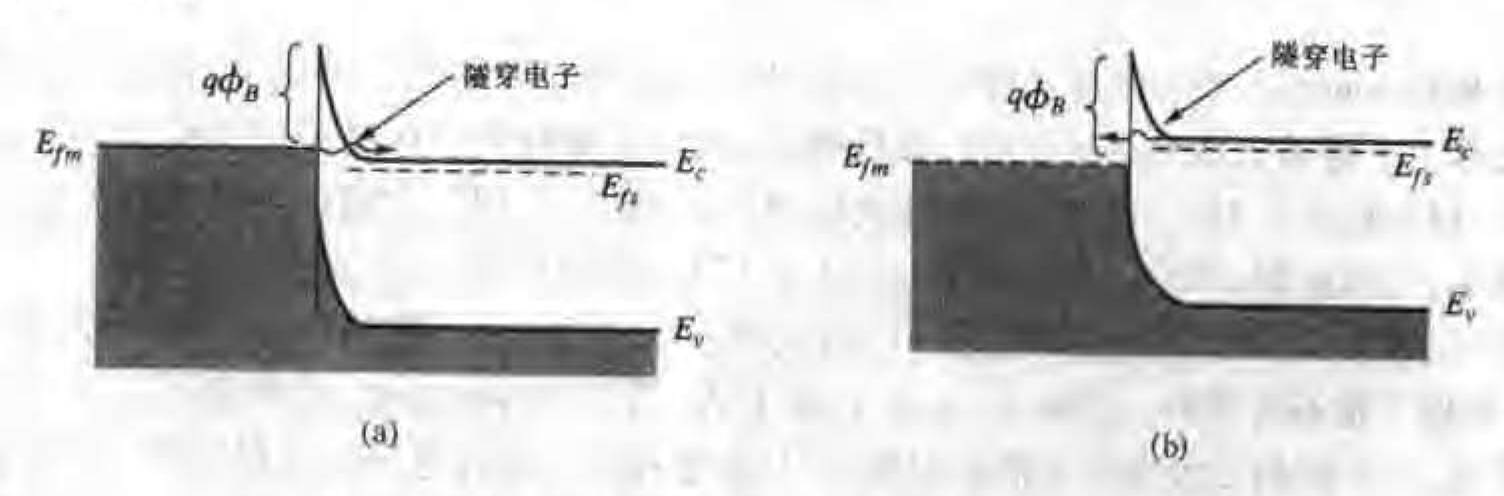


图3.12 空间电荷区很薄的金属-半导体势垒，电子可隧穿通过
(a)从金属到半导体的隧穿。(b)从半导体到金属的隧穿

现代器件中半导体区域的电导率远高于以前，因此欧姆接触必须具有更低的电阻才能保证其上几乎没有电压降。

3.4.2 Schottky 欧姆接触[†]

另一种获得欧姆接触的方法是使得结附近半导体中的多数载流子浓度远高于体内。如果金属-半导体达到平衡时，半导体表面出现多子积累而不是耗尽，这就形成了这个类型的欧姆接触。根据3.2节讨论的理想 Schottky 理论，这类情况对应于 n 型半导体的功函数高于金属的情况。这时，电子从金属流向半导体表面，在金属表面形成一正电荷层。[③] 图3.13(a)为相应的能带图，电荷与电场分布如图3.13(b)和3.13(c)所示。这些图与整流结对应的图3.4(b)和3.4(c)类似，两者重要的区别是，欧姆接触半导体中的电荷是自由电子组成的，而 Schottky 势垒半导体中的电荷是固定电荷(正的施主电荷)。Schottky 欧姆结的电荷、电场和电位分布都可以用与整流结类似的方法进行计算[5]，详细计算过程见习题3.11，结果总结如下。

为了计算半导体中空间电荷的分布，取金属中的 ϕ 作为电位的参考点。假定半导体表面区域的过剩电子(用 n' 表示)满足 Boltzmann 分布 $n' = n_s \exp(-q|\phi|/kT)$，其中 n_s 表示金属-半

③ 对 p 型半导体形成的欧姆接触，两种材料的功函数的相对大小刚好与 n 型半导体对应的情况相反，这样，半导体表面出现净的正电荷，因此在结附近形成空穴积累。

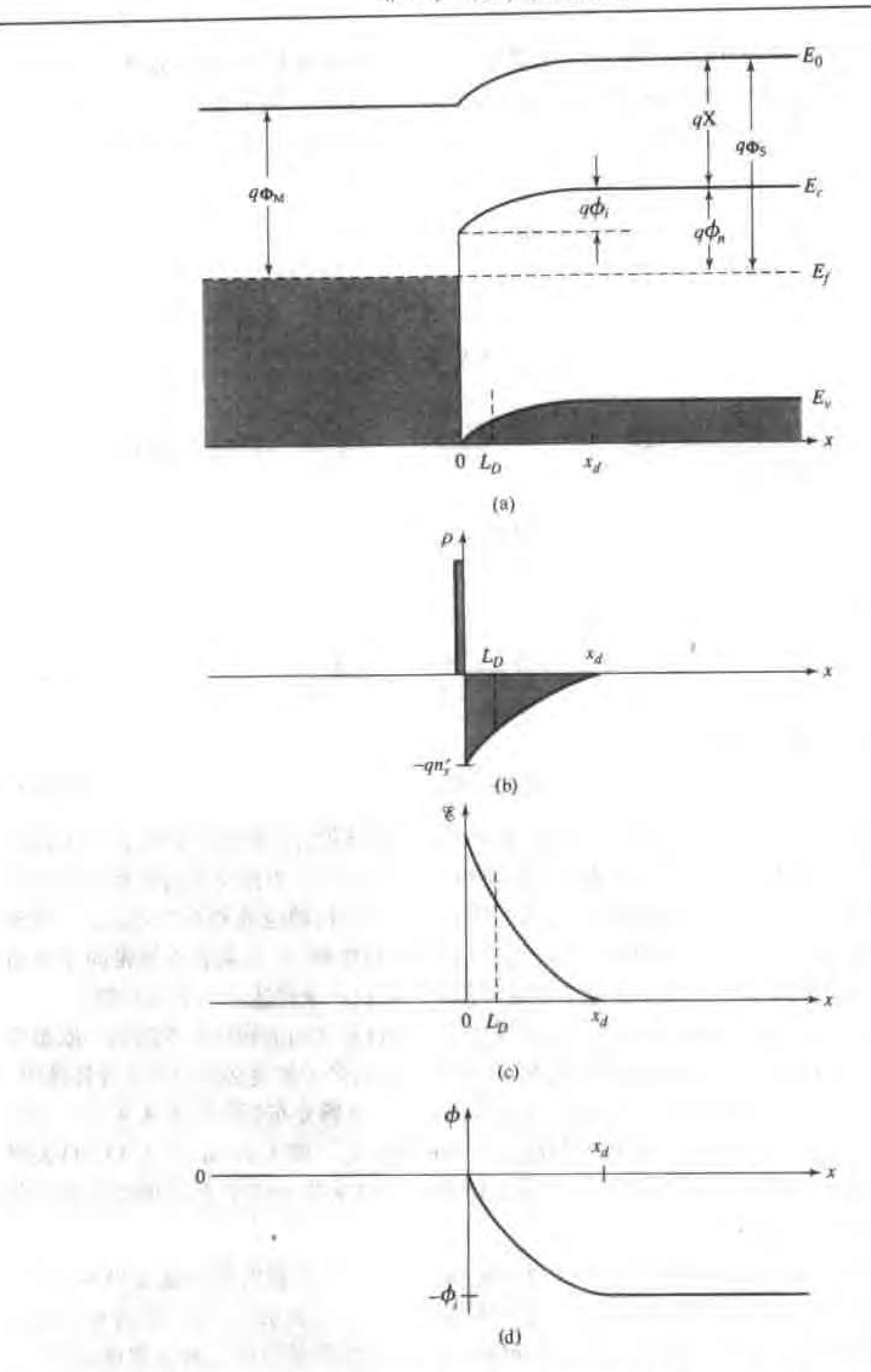


图 3.13　(a) 金属和 n 型半导体形成的理想的 Schottky 欧姆接触的平衡态能带图；(b) 理想的 Schottky 欧姆接触的电荷分布。金属表面以 delta 函数分布的正电荷与半导体表面的过剩电子浓度 $n'(x)$ 相平衡；(c) 理想的 Schottky 欧姆接触的电场分布；(d) 理想的 Schottky 欧姆接触的电位分布。Debye 长度 L_D 是描述电荷与电场分布的特征值

导体界面处半导体一侧的过剩电子浓度。如果半导体中的空间电荷密度用 qn' 表示(忽略施主离子对空间电荷的贡献),Poisson 方程可写成积分形式. 这一近似对空间电荷区的大部分区域都成立(除非电位与内建电势 ϕ_i 相差几个 kT/q 时)。解 Poisson 方程可得到空间电荷密度 $\rho(x)$

$$\rho(x) = -qn_s \Big/ \left(1 + \frac{x}{\sqrt{2}L_D}\right)^2 \tag{3.4.2}$$

其中

$$L_D = \left(\frac{\epsilon_s kT}{q^2 n_s}\right)^{1/2} \tag{3.4.3}$$

称做表面处的 Debye 长度。我们在结束这一部分的讨论后再详细介绍 Debye 长度。

电场随位置的变化为

$$\mathscr{E}_x = \frac{\sqrt{2}kT}{L_D q}\left(1 + \frac{x}{\sqrt{2}L_D}\right)^{-1} \tag{3.4.4}$$

半导体中的空间电荷层的宽度为

$$x_d = \sqrt{2}L_D\left[\exp\left(\frac{q|\phi_i|}{2kT}\right) - 1\right] \tag{3.4.5}$$

图 3.13(a)中的内建电势 ϕ_i 为

$$|\phi_i| = \phi_n - (\Phi_M - X) \tag{3.4.6}$$

其中半导体内的 $q\phi_n = (E_c - E_f)$。$\Phi_M - X = \phi_n$ 表示中性接触条件:内建电势为零,并且表面的自由电子浓度与体内相同。习题 3.12 表明当电流小于 $qn_0 v_{th}/4$ 时,中性接触可看做是欧姆接触,其中 n_0 为电子浓度,v_{th} 为电子的热运动速度。$\Phi_M - X \neq \phi_n$ 时,结电流相对于 $qn_0 v_{th}/4$ 受到因子 $\exp\{q[(\Phi_M - X) - \phi_n]/kT\}$ 的限制。因此,与 n 型材料接触时,如果表面能带向下弯曲则是欧姆接触,如果能带向上弯曲则对电子有阻碍作用。相反的条件适用于 p 型材料。

总之,总结一下欧姆接触的条件:多子在形成结的两种材料之间的转移不受阻碍。欧姆接触处通常存在内建电势。如果不存在隧穿势垒,欧姆接触处的多子浓度必须高于半导体体内

Debye 长度 空间电荷区的电荷密度(等式(3.4.2)),电场分布(等式(3.4.4))以及空间电荷区宽度(等式(3.4.5))的表达式中都包含 Debye 长度 L_D。图 3.13(a)至 3.13(d)表明 L_D 是边界处电场向空间延伸的定性度量。习题 3.13 的结果证实半导体中 50% 的空间电荷分布在距离表面 $\sqrt{2}L_D$ 长度内。

考虑更广义的范围,我们发现只要存在自由电荷时,Poisson 方程的解中就会出现 Debye 长度。如果电荷分布与上面不同,Debye 长度仍可用等式(3.4.3)表示,只是 n_s 变为相应的自由电荷浓度,我们将在第 4 章中举例说明这个问题。用 L_D 定性度量空间电荷区宽度,结论总是正确的。习题 3.14 给出了在特定区域内 L_D 与介电弛豫时间之间的关系。这个关系可以用自由载流子的两种输运机制之间的平衡从物理上得到解释:这两种输运机制是扩散(热运动引起的)和漂移(电场引起的)。

3.5　表面效应

3.2节考虑金属-半导体接触时，利用统一的Fermi能级描述平衡态时的能带图，这是完全理想化的，因为假定半导体和金属表面与体内的允许能态（能带）没有区别。而真实的情况更为复杂，下面将进一步推导，得到一个更加实际的物理模型。其中最重要的修正是考虑表面态效应。尽管这些效应修正了某些结论，但我们仍保留了前面的大部分概念。

3.5.1　表面态

先给出表面态的概念：表面态是存在于半导体表面的不同于体内的电子允许能态。这些能态有以下几个来源。首先考虑完全由晶格原子组成的清洁表面。因为晶体在表面的势场只分布在一边，电子只受到来自体内的束缚，如图3.14所示，所以在表面出现了额外的能态。我们认为在这些位置的电子特征能量不同于体内。④ 这种类型的表面态称为Tamm态或Schottky态以纪念首先研究它们的科学家[6,7]。对某一特定的半导体，Schottky－Tamm态的密度大约是表面原子面密度的量级，或约等于$N_0^{2/3}$（cm^{-2}），其中N_0是体原子密度（原子数/cm^3）。对硅来说，N_0为$5\times10^{22}cm^{-3}$（见表1.3），Schottky-Tamm态的密度约为$10^{15}cm^{-2}$。对金刚石晶格的研究表明[8]，这些态的分布在禁带内距价带边大约三分之一禁带宽度处有一峰值，如图3.15所示。尽管如此，关于这些态的能量分布目前仍不能很好地确定。

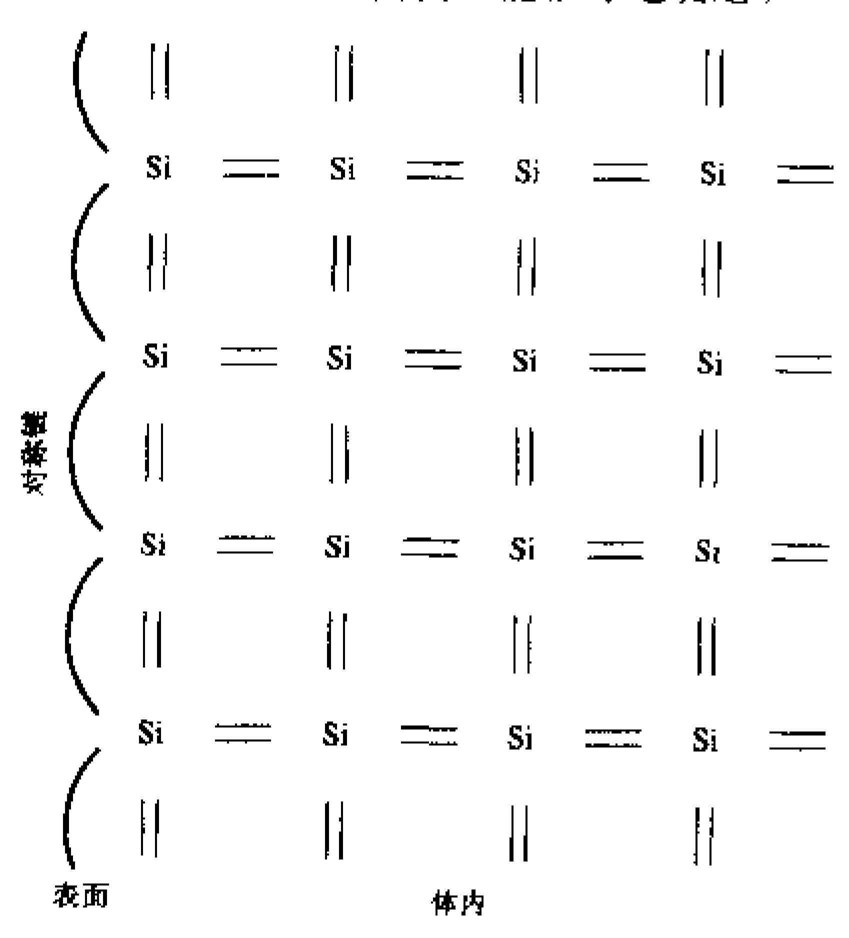


图3.14　硅晶体表面原子的成键图（直线表示共价电子对）。清洁的半导体表面的键是各向异性的，因此允许态能级不同于体内

表面态的其他来源还有表面处的外来原子和晶体的缺陷。其中在硅表面总是可以找到氧

④ 根据量子力学的观点，电子的波函数由于晶格势场的中断而受到微扰，因此表面的允许态能量不同于体内。

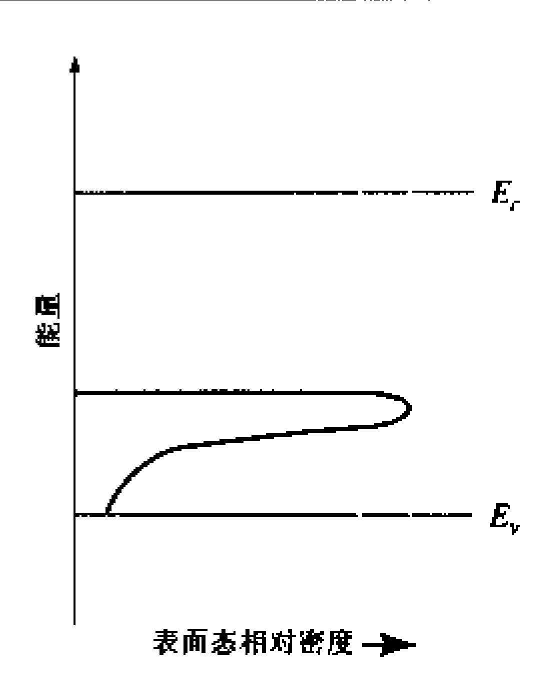


图 3.15 金刚石晶格中的 Tamm - Schottky 态的近似分布[8]。高于 E_v 三分之一禁带宽度处有一尖峰

就是一个例子。氧形成的表面态可以在很宽的能量范围内分布,取决于它与硅原子的成键的特性。在任何加工后的硅表面总是存在金属络合物、氢氧基离子等,这些因素的总效应是在任何能量下表面电子态密度都不为零,尽管在某些能量会出现明显的峰值。

除了随能量变化外,表面态分布还随类型变化,可根据平衡态时的带电状态来划分。如果有电子占据时呈电中性,没有电子占据时带正电,这类表面态称为施主态;有电子占据时带负电,没有电子占据时呈电中性,这类表面态称为受主态,这与 1.1 节中讨论的半导体杂质原子的分类类似。

另一种表面态的分类(有时称为界面态)是根据固体间的实际界面特性来划分的。在原子尺度上(小于一个纳米),这样的界面由厚度为几个到几十个原子层的过渡材料和杂质组成,而不是突变的。在过渡区内,某些表面态的物理特性接近于体半导体,这些表面态与体内的能态保持热平衡,即使在偏压变化相当快时。因为占据这些态的电子能很快达到平衡态,所以这些态称为表面快态。与之相反的就是所谓的慢态,这些态位于距离半导体体内较远的过渡层,需要较长的时间才能与体内电子达到热平衡。虽然对界面态类型这样划分比较模糊,但一般的做法是取 1kHz 对应的响应时间作为界限。

3.5.2 金属-半导体接触的表面效应†

如果存在表面态,需要修正前面提到的金属-半导体接触理论。如果半导体表面态不是中性的,或者形成接触时带电状态发生了变化,在 Schottky 理论中得到的电荷分布将不再正确。文献[9]已经给出了具有界面态的金属-半导体系统理论。本书将不详细讨论,仅给出结论以及实验验证。

为了解释界面效应,金属-半导体接触可看成两种晶体之间包含一层过渡区的结构。过渡区厚度为几个到十几个原子尺度,这一层包含杂质原子和引入的界面态。过渡层很薄,不会阻碍电子转移(通过隧穿),但是在它上面有电压降。假定表面态引起的允许电子态分布在过渡层和半导体之间的界面上,能带结构和表面态分布如图 3.16 所示。图中假定表面态为受主型,态密度为 $D_s(\mathrm{cm^{-2}eV^{-1}})$。注意,图 3.16 中过渡层厚度为 δ,其上的电压降为 Δ 伏特。因为过渡层厚度不足以阻挡电子隧穿,所以金属-半导体势垒高度仍为金属 Fermi 能级与半导体界

面处导带之间的能量差。即使不存在金属，表面层也会影响到半导体的特性，如图 3.17 所示。当存在受主表面态时，如图 3.17 所示，n 型半导体靠近表面处的电子耗尽，受主表面态带负电。

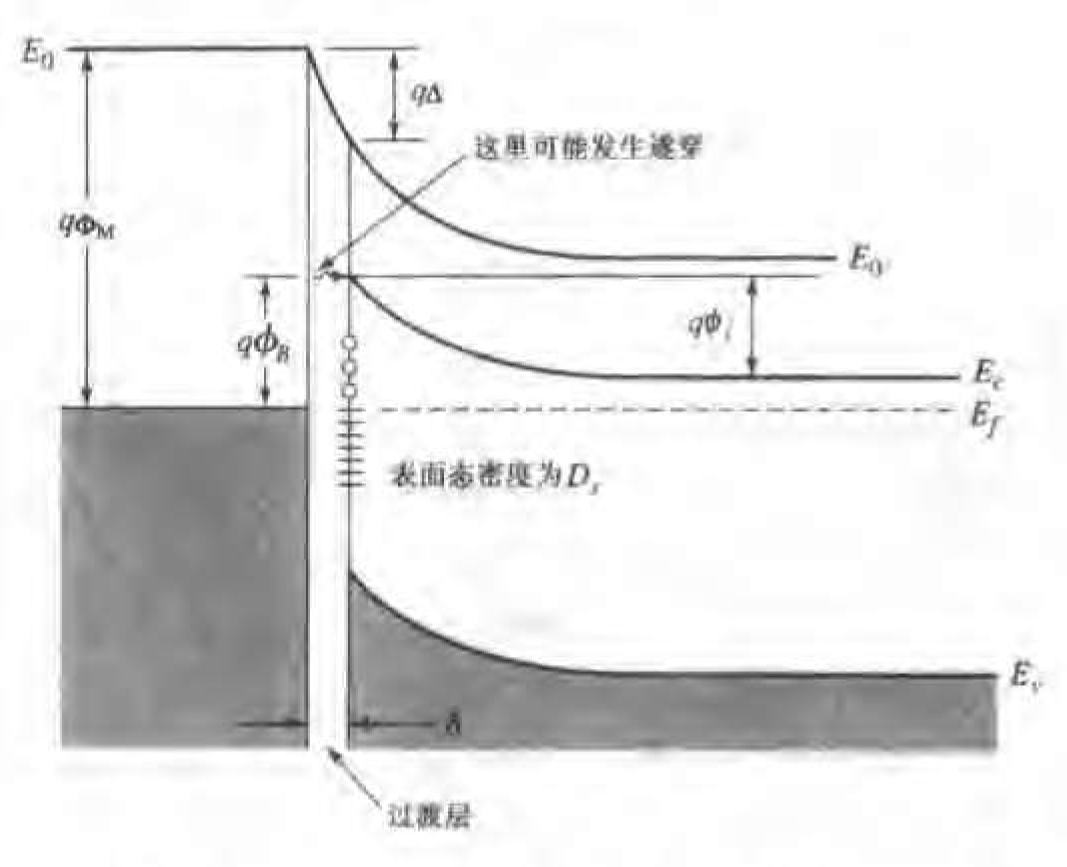


图 3.16　根据 Cowley 和 Sze[9] 模型的金属-半导体接触的能带图。模型考虑了厚度为 δ 的薄过渡层，平衡时过渡层上的电压降为 Δ。假定受主型表面态密度为 D_s

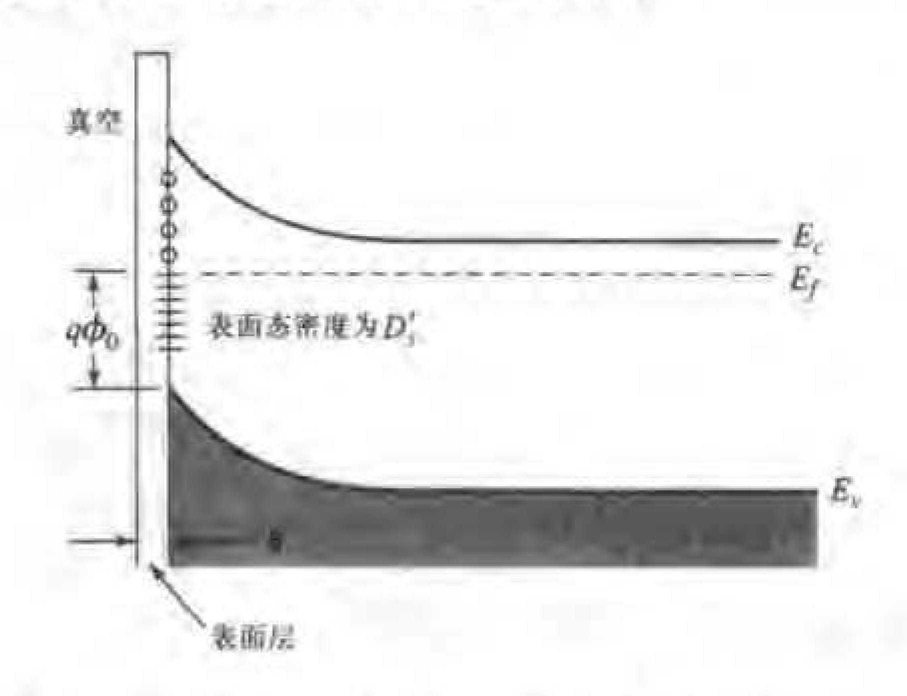


图 3.17　包含受主型表面态的半导体表面的能带图。由于表面态上的电荷转移形成了表面耗尽区

考虑图 3.17 所示的具有表面态的半导体阻挡型接触的形成，它的能带图与图 3.16 类似。根据 Schottky 理论，如果金属的功函数 Φ_M 大于半导体的功函数 Φ_S，电子将从半导体流向金属，形成阻挡型接触，电子的转移使半导体导带边向上弯曲。这样，在图 3.17 的半导体中，位于 E_f 之上的部分表面态不再被电子占据。D_s 越大，接触处 E_c 向上弯曲所需要移去的电荷越多。如果 D_s 很大，Fermi 能级在半导体表面的微小移动就需要转移足够多的电荷才能达到平衡。在这种情况下，我们说 Fermi 能级被高的态密度钉扎。注意 Fermi 能级钉扎并不只是与受

主型表面态有关。只要态密度非常大,小的 Fermi 能量的变化就能导致大量的电荷转移,所以 Fermi 能量附近的任何高密度电子态都会引起 Fermi 能级的钉扎。

当 Fermi 能级被钉扎,势垒高度 $q\phi_B$ 变为[19]

$$q\phi_B = (E_g - q\phi_0) \tag{3.5.1}$$

其中 $q\phi_0$ 为没有金属电极时半导体表面的$(E_f - E_v)$,如图 3.17 所示。当 D_s 接近于零时,势垒高度 $q\phi_B$ 接近于等式(3.2.1)中基本 Schottky 理论计算出的势垒高度,我们重复写在这里

$$q\phi_B = q(\Phi_M - X) \tag{3.5.2}$$

其中金属-半导体势垒高度是用等式(3.5.1)还是用等式(3.5.2)计算或者等于两者之间的某个值,取决于 Fermi 能级附近 D_s 的大小以及界面的特性,比如它的精确厚度和介电常数等[9]。

实际上,大部分重要的半导体 Schottky 势垒高度用等式(3.5.1)计算比用等式(3.5.2)更为准确,并且势垒高度与金属功函数的依赖关系很弱。对硅、锗,尤其是砷化镓和其他Ⅲ-Ⅴ族半导体材料都是这样的。实验得到的硅、锗、砷化镓和磷化镓的 $q\phi_0$ 大约为 $E_g/3$。因此根据等式(3.5.1),势垒高度 $q\phi_B$ 的典型值接近带隙的三分之二,对硅大约为 0.75eV。Fermi 能级被钉扎在这个能量上,这一现象是高态密度材料如金刚石结构晶体的典型特征。不同金属与 n 型和 p 型硅晶体的接触势垒的测量值如表 3.1 所示。

表 3.1 硅的 Schottky 势垒高度[10](硅的 $qX = 4.05$eV)

导电类型	金属	$q\Phi_M$(eV)	$q\phi_B$(eV)
n	Al	4.1	0.69
p	Al		0.38
n	Pt	5.3	0.85
p	Pt	—	0.25
n	W	4.5	0.65
n	Au	4.75	0.79
p	Au	—	0.25

3.6 金属-半导体器件:Schottky 二极管

金属-半导体势垒器件,也就是通常所说的 Schottky 二极管,在工业上应用很广泛。最重要的应用领域是数字逻辑电路,即二进制电路。这些电路中 Schottky 二极管通常用做快速开关,可制作在小尺寸的集成电路芯片上。人们对 Schottky 二极管功率整流器的兴趣越来越大,因为与金属接触的大面积器件具有良好的热传导,可以在大电流下工作。与第 4 章和第 5 章介绍的扩散形成的 pn 结二极管相比,流过大电流时 Schottky 二极管结上的电压低一些。

在微波范围,Schottky 二极管也可用做可变电容器,并且能有效地工作。此时,Schottky 二极管工作在负偏压下,由 3.2 节知,利用电压的变化可调节耗尽层的宽度和电容。

利用耗尽层宽度调制效应的另一个应用是Schottky势垒场效应晶体管[5]，也称为金属-半导体场效应晶体管(MESFET)。如图3.18所示，这种器件中的肖特基结作为输入控制电极(或栅)，两个欧姆接触作为输出端(称为源极和漏极)。栅电极下的导电沟道的截面积发生变化时，输出电流也将发生变化。这种结构是结型场效应晶体管(JFET)——使用反偏pn结作为控制电极的一种放大器件——的一种特殊形式。用Schottky势垒代替pn结尤其有用，因为制备工艺比pn结简单。使用Schottky势垒的MESFET主要应用于制备高频GaAs器件。我们将在第4章介绍pn结的工作原理后，详细讨论JFET和MESFET的工作原理。

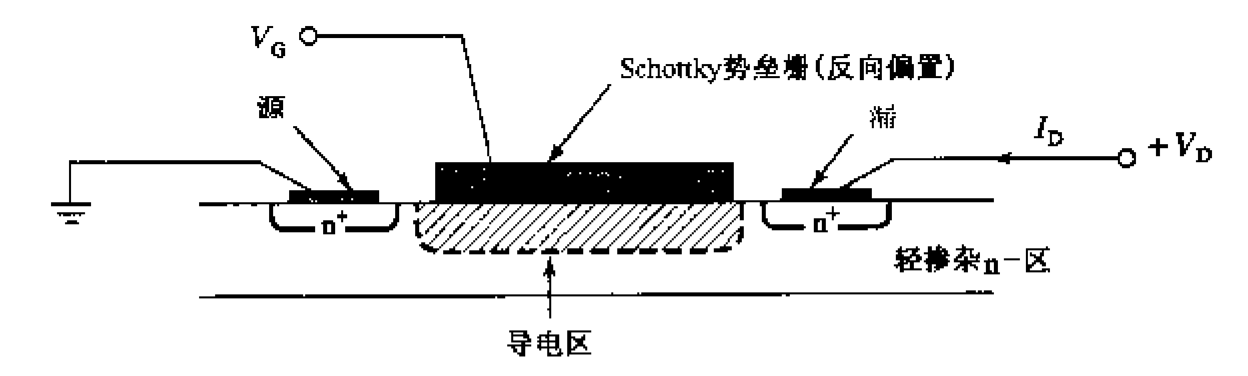


图3.18 Schottky势垒场效应晶体管。栅电压V_G通过控制耗尽区的尺寸来调节从漏到源的电流I_D，也就是说改变了I_D流过的导电区域的截面积。源极和漏极制作在高掺杂材料上，为欧姆接触

集成电路中的Schottky二极管

硅数字集成电路中制备Schottky二极管的方法非常简单，这个方法来源于两个与制备Schottky势垒无关的偶然事件：采用高阻n型硅制作npn双极晶体管和蒸发金属铝形成集成电路中的"互连线"。如果硅表面彻底清洗，铝在轻掺杂n型硅表面将形成阻挡型Schottky结。正如3.4节所描述的，只要硅的掺杂足够低，电子将不能隧穿通过势垒。实际上，这就要求Schottky二极管中硅的掺杂浓度要小于$10^{17}cm^{-3}$(习题3.7)。

n型硅与铝的势垒高度大约为0.7eV，在真空中硅表面淀积铝形成的二极管的特性与理论预测值吻合很好，如图3.10所示。然而，电力线密度在拐角处很高，因此反向击穿不是硬击穿，并且击穿电压较低(-15V)。为提高其反向特性，发展了几种保护技术，如扩散保护环(图3.19(a))或场板(图3.19(b))。然而，由于保护电路制作工艺复杂，除非特别需要，一般都不采用这些技术。可用难熔金属制备高势垒的性能优越的Schottky二极管，尤其是铂经常被采用。因为难熔金属的熔点较高，很难蒸发，所以通常用溅射来淀积[6]。因为溅射的淀积速率高，熔点低的金属也经常采用这种方法。如果制备工艺中引入溅射，那么硅表面可用高能离子轰击来彻底清洁(溅射剥蚀)。表面剥蚀后，溅射淀积的铂将形成性能优越的Schottky二极管。尽管Schottky二极管也可制作在p型硅上，但由于势垒高度较低(大约为带隙的三分之一或0.36V)，因此电学特性相应较差。

数字逻辑电路中Schottky二极管称为箝位二极管，因为它们固定或箝制了晶体管中结上

[5] 因为这是本书第一次遇到晶体管(Transistor)这个词，需要指出的是这个词是由传输(Transfer)和电阻(Resistor)两个词合并得到的。晶体管是由W. Shockley和他在贝尔实验室的同事共同命名的，第6章和第7章将要讨论的结型场效应晶体管和双极晶体管都是在贝尔实验室发明的。

[6] 溅射是指用高能气态离子(通常是氩离子)轰击源或靶淀积薄膜材料的一种技术。

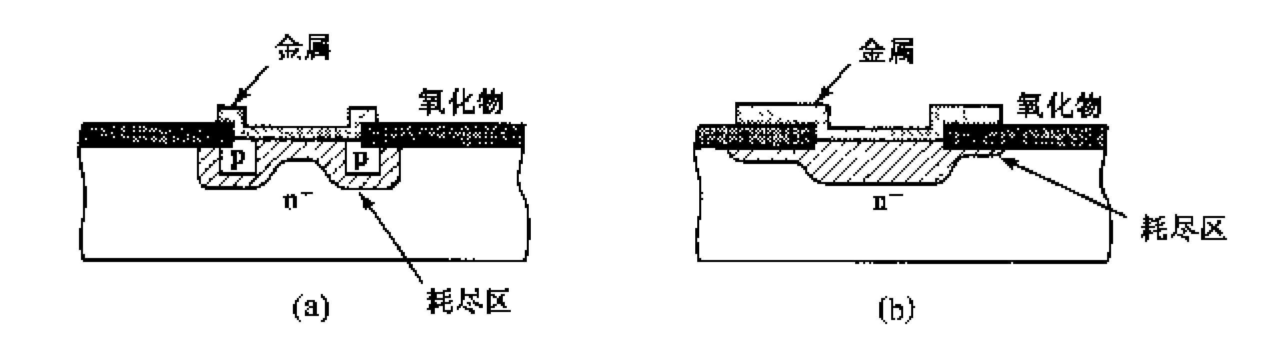


图 3.19 改善 Schottky 二极管性能的特殊加工工艺的横截面图。(a)扩散形成 p 型保护环使电场分布均匀,消除了结的边缘和拐角处的击穿;(b)金属场板是另一种达到同样效果的方法

的电压,提高了电路性能。我们将在第6章中具体解释,这里只说明如果箝位二极管放置在开关晶体管的集电极和基极之间,将能保证集电结不会正偏,从而提高了数字逻辑电路的速度。Schottky 二极管是这种应用的最佳选择。

为了说明这个问题,我们先介绍开启电压的概念。开启电压是指二极管开启时的外加正向电压,即二极管中有一定电流流过时的电压。金属-半导体势垒的电流、电压方程(等式(3.3.17)和(3.3.19))表明电流随电压连续变化,没有可用来确定开启电压的电流突变特性。然而,从工程的角度看,在线性坐标下绘制 Schottky 二极管的电流与电压关系时,对应于导通,有明显的阈值(图 3.20)。线性电流坐标限制了描述电流的范围。电流与电压的强烈依赖关系可以用两段直线对实验数据进行相当好地拟合:一条几乎与直线 $J=0$ 重合,另一条几乎与电压轴垂直。垂直线与电压轴的截距被定义为开启电压 V_o。在设计数字电路时,任意导通二极管上电压降近似等于 V_o。根据等式(3.3.17),在给定的正向电流密度 J_F 下,V_o 为

$$V_o = \frac{nkT}{q}\ln\left(\frac{J_F}{J'_S}+1\right) \tag{3.6.1}$$

这样,从设计者的观点来看,V_o 与 J'_S 的关系最密切。n 型硅上制备的铝 Schottky 二极管的工作电流在毫安范围时,V_o 大约为 450mV。这一数值比 pn 结二极管的 V_o(第5章)约小 200mV,所

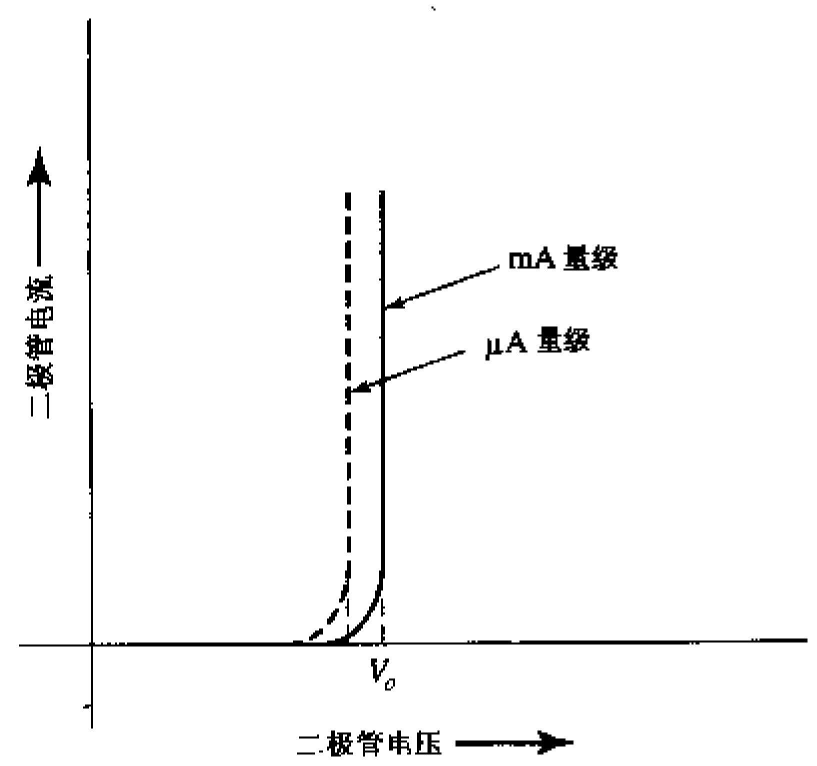


图 3.20 线性坐标下的 Schottky 二极管的电流电压关系,V_o 是二极管的"开启电压"

以 Schottky 二极管与 pn 结二极管并联时，正向偏压将不会使 pn 结二极管导通。Schottky 箝位集电结时，双极晶体管不会进入使开关速度变慢的饱和区，因此，使用 Schottky 箝位二极管的数字电路比没有箝位的电路要快几个纳秒。

使用 Schottky 二极管在电学性能上付出的惟一代价是反偏的 Schottky 二极管使电路的负载电容有少量的增加。箝位晶体管比未箝位晶体管占用硅芯片的面积略微大一点，但却对集成电路带来了很多优点。当然，采用 Schottky 工艺会使成品率降低，因为金属化和表面处理工艺的要求比起不使用 Schottky 二极管的电路来说更加严格。

小结

热平衡状态下的所有电子可用统一的 Fermi 能量表征，因此，Fermi 能量是电子允许态随位置变化的参考值，在绘制非均匀材料以及密切接触材料系统的热平衡态的能带图时尤其有用。如果 Fermi 能量不同（因此彼此不处于热平衡）的两种材料结合在一起时，电子将从 Fermi 能量高的材料流向 Fermi 能量低的材料。在体内能态分布与表面一致的理想情况下，并且半导体材料均匀，利用上述结论，可得到 Schottky 金属-半导体接触的基础理论。这一理论指出，如果金属功函数 Φ_M 大于半导体功函数 Φ_S，则金属与 n 型半导体形成阻挡结或整流结，而 Φ_S 大于 Φ_M 时则形成欧姆结。对金属和 p 型材料形成的接触，将这个结论反过来即可。理想 Schottky 接触的能带如图 3.21 所示。如果 Schottky 势垒处的空间电荷区足够薄，电子可隧穿通过（发生在半导体重掺杂时），形成欧姆接触。

为了得到金属-半导体接触的理论，需要求解 Poisson 方程并利用热平衡时的能带图。耗尽近似和准平衡等假定可以简化理论推导。基本的 Schottky 理论成功地预测了许多可观察到的特性，其中最主要的是 Schottky 势垒的电流电压关系和反偏电容特性。用同样的分析方法研究在低阻衬底上的薄的高阻区上形成的接触，可得到 Mott 势垒的有关结论。从 Schottky 欧姆接触的分析中引入了 Debye 长度的概念，Debye 长度是在自由电荷浓度较高区域电场向半导体区域延伸程度的度量。尽管基本的 Schottky 理论为金属-半导体接触提供了许多有用的结论，但不能准确求出与硅接触的势垒高度。该理论主要的缺陷在于对表面效应的处理。表面态源于表面周期性晶格的中止、缺陷和表面杂质。金属与真实的硅表面形成阻挡结的理论基于存在组分不确定的但具有特定电学特性的薄的过渡层的假定。金属-半导体接触形成的 Schottky 二极管应用非常广泛。Schottky 二极管很容易用标准的硅平面工艺制备，可以应用在集成电路中。

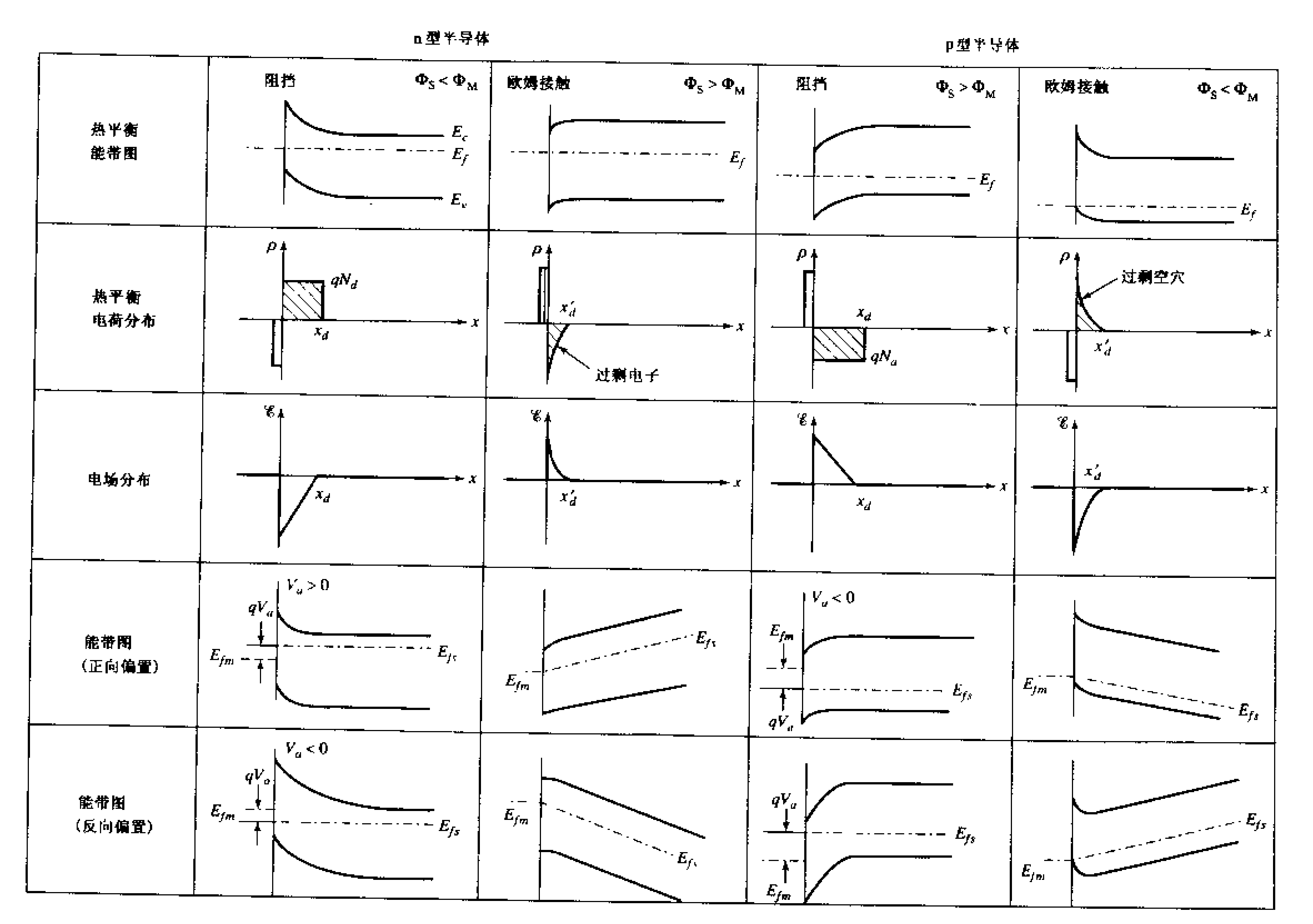


图 3.21 理想的金属—半导体 Schottky 二极管的特性图

参考文献

1. W. Schottky, *Naturwissenschaften* **26**, 843 (1938).
2. (a) H. A. Bethe, *Theory of the Boundary Layer of Crystal Rectifiers*, MIT Radiation Laboratory Report 43-12 (1943).
   (b) S. M. Sze, *Physics of Semiconductor Devices, 2nd Edition*, Wiley-Interscience, New York (1981), pp. 255-258.
3. H. K. Henisch, *Rectifying Semiconductor Contacts*, Oxford at the Clarendon Press (1957), p. 172.
4. N. F. Mott, *Proceedings Cambridge Philosophical Society* **34**, 568 (1938).
5. A. Rose, *Concepts in Photoconductivity and Allied Problems*, Wiley-Interscience, New York (1963).
6. I. Tamm, *Phys. Z. Sowjetunion* **1**, 733 (1933).
7. W Shockley, *Phys. Rev.* **56**, 317 (1939).
8. D. Pugh, *Phys. Rev. Lett.* **12**, 390 (1964).
9. A. M. Cowley and S. M. Sze, *J. Appl. Phys.* **36**, 3212 (1965).
10. S. M. Sze, *Physics of Semiconductor Devices. 2nd Edition*, Wiley-Interscience. New York (1981)

参考书

S. M. Sze, *Physics of Semiconductor Devices*, 2nd Edition Wiley-Interscience, New York (1981).

A. G. Milnes, *Semiconductor Devices and Integrated Electronics*, Van Nostrand Reinhold, New York (1980).

习题

3.1* 已知硅中的施主浓度从$5\times10^{16}cm^{-3}$突变到$8\times10^{15}cm^{-3}$ 画出与图1.11类似的能带图

(a)求出硅晶体中这两个区域的功函数。

(b)求出硅晶体中这两个区域之间的电位差。

3.2 图P3.2(a)是相同的坐标系统中金属和半导体的能带图。图P3.2(b)~(d)给出了热平衡条件下三种可能的金属-半导体接触的能带图

(a)解释为什么每一个图都是错的

(b)利用基本的Schottky理论,画出正确的金属-半导体能带图

3.3 根据简单的Schottky理论,考虑金属-半导体接触:

(a)画出铜(功函数=4.5eV)与硅(功函数=4.25eV)接触的理论能带图.

(b)如果光照射在结上产生电子空穴对:

  (i)如果结与电路连接,器件中的电流将如何流动?

  (ii)通过结上测出的最大电压是多少(电流输出为零)?

(c)画出铜与功函数为4.9eV的硅接触时的能带图?

(d)比较(a)和(c)中的金属-半导体系统的电学特性

3.4 图P3.4给出了接触面积相同的不同金属与硅接触的数据。如果使用Schottky理论,试问哪种金属可能具有较高的功函数?哪组数据是在电阻率为1Ω·cm的硅上测量得到的?哪组数据是在电阻率为5Ω·cm的硅上测量得到的?判断你的答案是否正确,并解释为什么用“可能”这个词。(考虑3.5节)

3.5*(a)铂(功函数=5.3eV)与掺杂浓度$N_d=10^{16}cm^{-3}$的硅形成理想Schottky势垒。计算300K时直流偏置为零的小信号电容。已知Schottky二极管面积为$10^{-5}cm^2$。

(b)当电容比零偏电容减小了25%时,计算此时的反向偏压。

3.6*(a)计算电子势垒取极大值的位置x_m[图3.8中的$E_2(x)$],并证明等式(3.2.13)。

(b)当外加电场为$10^5V\cdot cm^{-1}$时,计算x_m和$q\Delta\phi$。

3.7* 考虑铝与均匀掺杂的n型硅形成的Schottky势垒,已知掺杂浓度为N_d,势垒高度$q\phi_B=0.65eV$

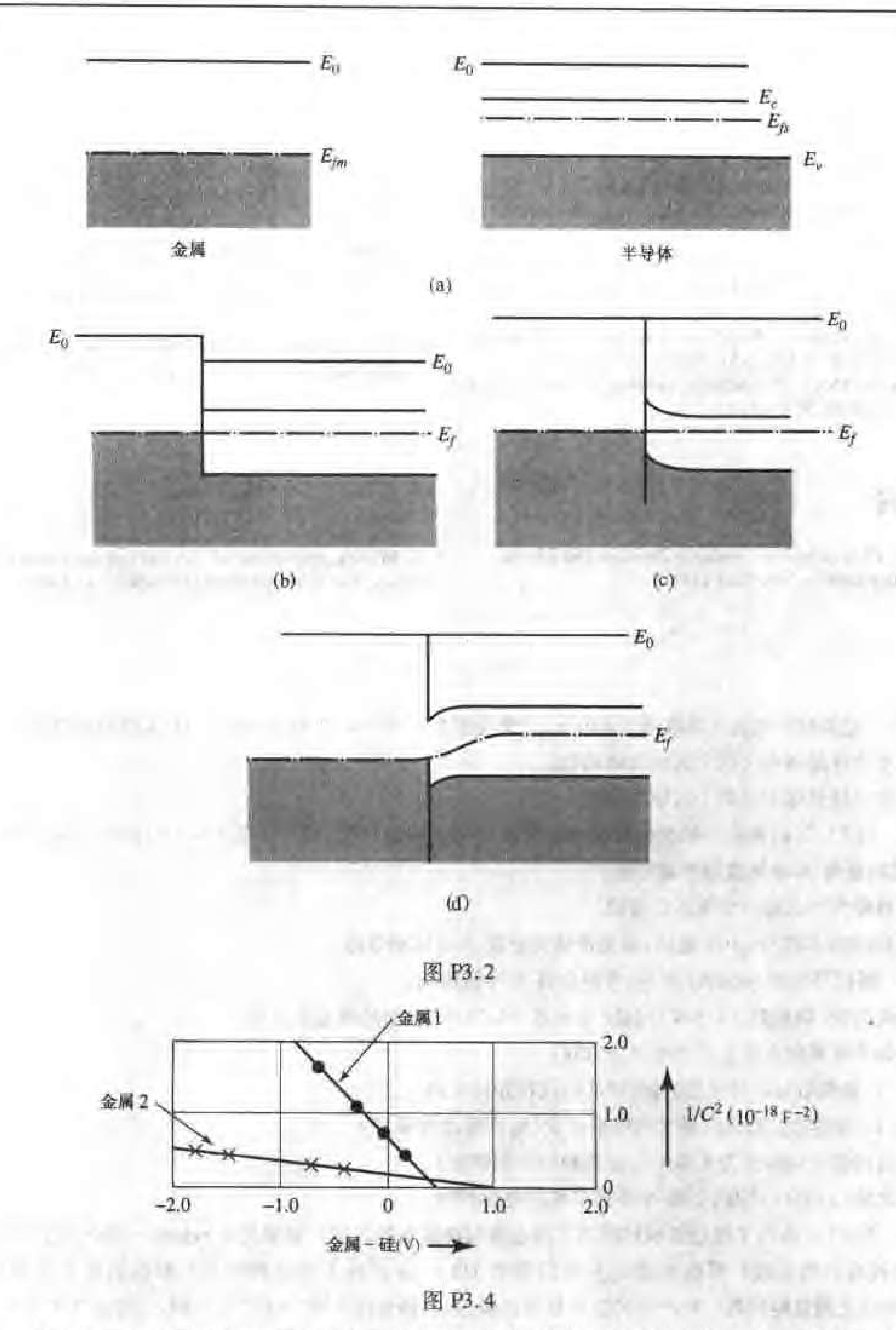


图 P3.2

图 P3.4

在反向偏置下,如果势垒足够薄,电子将从金属隧穿通过势垒,结中将流过很大的电流。假定金属中的 Fermi 能量与半导体内 10nm 处的导带边(E_c)相等时,开始发生有效的隧穿。

(a)如果结上的总压降($\phi_i - V_a$)=5V 时上述条件成立,请问 N_d的最大值为多少?

(b)这将对 Schottky 箝位电路中的外延层的电阻率有什么限制?

(c)画出发生有效隧穿时的能带图。

3.8[†] 推导等式(3.3.13)。

3.9[+]　在小的正向偏压下考虑等式(3.3.16)。证明等式(3.3.17)可用$(1-V_a/\phi_i)^{1/2}=\exp\left[\frac{1}{2}\ln(1-V_a/\phi_i)\right]$及$J_s$的近似表达式来推导。这种方法推导出的等式(3.3.17)中的$n=(1+kT/2q\phi_i)$，一般来讲比测量值小。但是，其他效应如势垒降低对等式(3.3.17)中n值也会有影响，导致n值比本题推导出来的值稍高一些。

3.10　使用理想的二极管关系(等式(3.3.6))，在线性坐标系中画出$I-V_a$曲线，已知：

(a)$I_0=1\text{pA}$，$T=150\text{K}$。

(b)$I_0=1\text{nA}$，$T=300\text{K}$。

(c)$I_0=1\mu\text{A}$，$T=450\text{K}$。

(d)考虑3.6节中的内容，对各条$I-V_a$曲线求出合适的开启电压V_o。

3.11[*]　利用3.4节的公式可以表示金属和n型半导体形成的Schottky欧姆接触的空间电荷并建立Poisson方程，方程的形式为$d^2\phi/dx^2=K\exp(\phi/V_t)$，其中$V_t=kT/q$。利用等式$d^2\phi/dx^2=\mathscr{E}\,d\mathscr{E}/d\phi$，很容易将电位$\phi$与位置$x$的函数关系转化为电场$\mathscr{E}$与电位$\phi$的函数关系。最后可推导出等式$\mathscr{E}=(2n_s kT/\epsilon_s)^{1/2}\exp(\phi/2V_t)$。

(a)根据前面描述的思路推导本题中的公式。

(b)根据相同的分析方法推导等式(3.4.2)、(3.4.4)和(3.4.5)。

3.12[+]　画出3.4节中描述的"中性"接触的能带图。考虑习题1.13中的结果，将半导体中自由电子的热运动形成的电流表示为$qn_0v_{th}/4$，其中n_0为电子浓度，v_{th}为电子热运动速度。如果从金属流向半导体的电流小于这个值，则结对电流没有限制，可视为欧姆结。

(a)如果半导体中的电场小于$v_{th}/4\mu_n$时，证明是欧姆接触。

(b)计算掺杂浓度$N_d=10^{16}\text{cm}^{-3}$的半导体中性接触的欧姆接触限制电流，取$A=10^{-5}\text{cm}^2$，$v_{th}=10^7\text{cm/s}$

(c)如果能带弯曲并且$q(\Phi_M-X-\phi_n)=0.65\text{eV}$，计算欧姆接触限制电流是多少？

3.13[+]　利用等式(3.4.2)和(3.4.5)证明Schottky欧姆接触的半导体中，50%的空间电荷位于距离表面$\sqrt{2}L_D$的范围内。

3.14　证明介电弛豫时间$\tau_r=\epsilon_s/\sigma$(见习题1.12)与Debye长度L_D之间的关系为$L_D=(D\tau_r)^{1/2}$，其中D为材料的扩散系数。

3.15　假定基本的Schottky理论成立，

(a)画出平衡态下金属与p型硅之间欧姆接触的能带图；

(b)画出反向偏压为2V时金属与p型硅之间阻挡结的能带图。

3.16[*]　集成电路中的Schottky势垒二极管和欧姆接触都是在硅上淀积金属获得的。已知金属的功函数为4.5eV，考虑理想的Schottky特性，计算每一种接触允许的硅掺杂范围。要求考虑n型和p型硅，并对给定掺杂条件下的集成电路加工工艺进行评述。

3.17[+]　与硅接触形成的反偏Schottky二极管可用做收音机中的调谐元件(550～1650kHz)。为了操作方便，希望调谐电路的共振频率($1/2\pi\sqrt{LC}$)随加在电路上的0～5V的直流电压线性变化。如果调谐电感为2mH，可以很容易计算出两个极端电压处的电容值分别为41.8pF和4.65pF。假定二极管面积为10^{-3}cm^2，计算所需的掺杂浓度N_d(将计算结果画在半对数坐标系中)。提示：根据给定的信息有

$$\frac{df}{dV}=-\frac{1}{4\pi\sqrt{LC}}\frac{1}{C}\frac{dC}{dV}=0.22\ \text{MHz/V}$$

利用等式(3.2.10)及$C=A\epsilon_s/x_d$求解$N(x_d)$

第4章　pn　结

从第3章中知，热平衡条件下的电子系统具有统一的Fermi能级。这个原理最初用来推导两种不同掺杂浓度的半导体的能带图。电子从Fermi能级较高的区域流向Fermi能级较低的区域，达到热平衡状态。电荷的迁移将会形成一个势垒，阻止电子的进一步流动，同时，势垒上的电压降增加到恰使两种半导体的Fermi能级达到一致。以上这些概念就是分析金属-半导体接触的理论基础。

本章将考虑具有不同掺杂浓度的单晶半导体材料中类似的问题。如果采用在器件分析中常用的两个重要近似，将会有利于问题的分析。一个是耗尽近似，这在第3章中就已提及；另一个是准中性近似，也可以简化计算。准中性近似集中讨论器件给定区域的最重要的物理特性，使复杂问题易于处理。它适用于讨论掺杂浓度缓慢变化的半导体区域，而耗尽近似主要用于分析密切接触的p型和n型半导体。

下面将详细讨论pn结的过渡区以及与之相关的势垒分布，然后考虑反向偏压对过渡区的影响，这将导致电容效应，接下来将这个概念引申到两种不同半导体材料系统中。我们还将讨论两种重要的击穿机制对外加反向偏压的限制，因为大部分半导体器件包含一个或多个pn结。这些结论可直接用于实际器件的分析。本章的最后一节将介绍利用反偏pn结概念的结型场效应晶体管。

本章重点讨论平衡状态和反向偏置下的pn结。我们认为可以忽略pn结反向电流，除非结上的反偏电压大到足以产生击穿。关于pn结正向电流将在第5章讨论。

4.1　缓变杂质分布

本节考虑半导体掺杂浓度在空间任意分布的平衡问题，如图4.1(a)所示。假设最初材料中的任意一点的多数载流子浓度等于掺杂浓度——这是一个非平衡条件。然后，我们来研究系统趋于热平衡的途径。由等式(1.2.17)知可动载流子的浓度梯度将导致载流子从高掺杂区向低掺杂区扩散。载流子离开原位置将留下未被补偿的、带相反电荷的杂质离子。这种正负电荷的分离产生了一个阻止扩散流的电场。当载流子浓度梯度造成的扩散电流和电荷分离产生的电场引起的方向相反的漂移电流在数值上相等时，就达到了平衡。这样平衡时可动载流子的分布并不恰好与固定杂质分布一致，并且内建电场会阻止这两种电荷的进一步分离。由这种机制产生的空间电荷只是掺杂浓度[1]的一小部分，但导致的电场却对器件的性能有严重的影响。

下面讨论内建电场如何影响能带图。因为系统处于热平衡，整个系统的Fermi能级一致。

---

[1] 这一结论将在4.2小节的例题中进一步考虑。

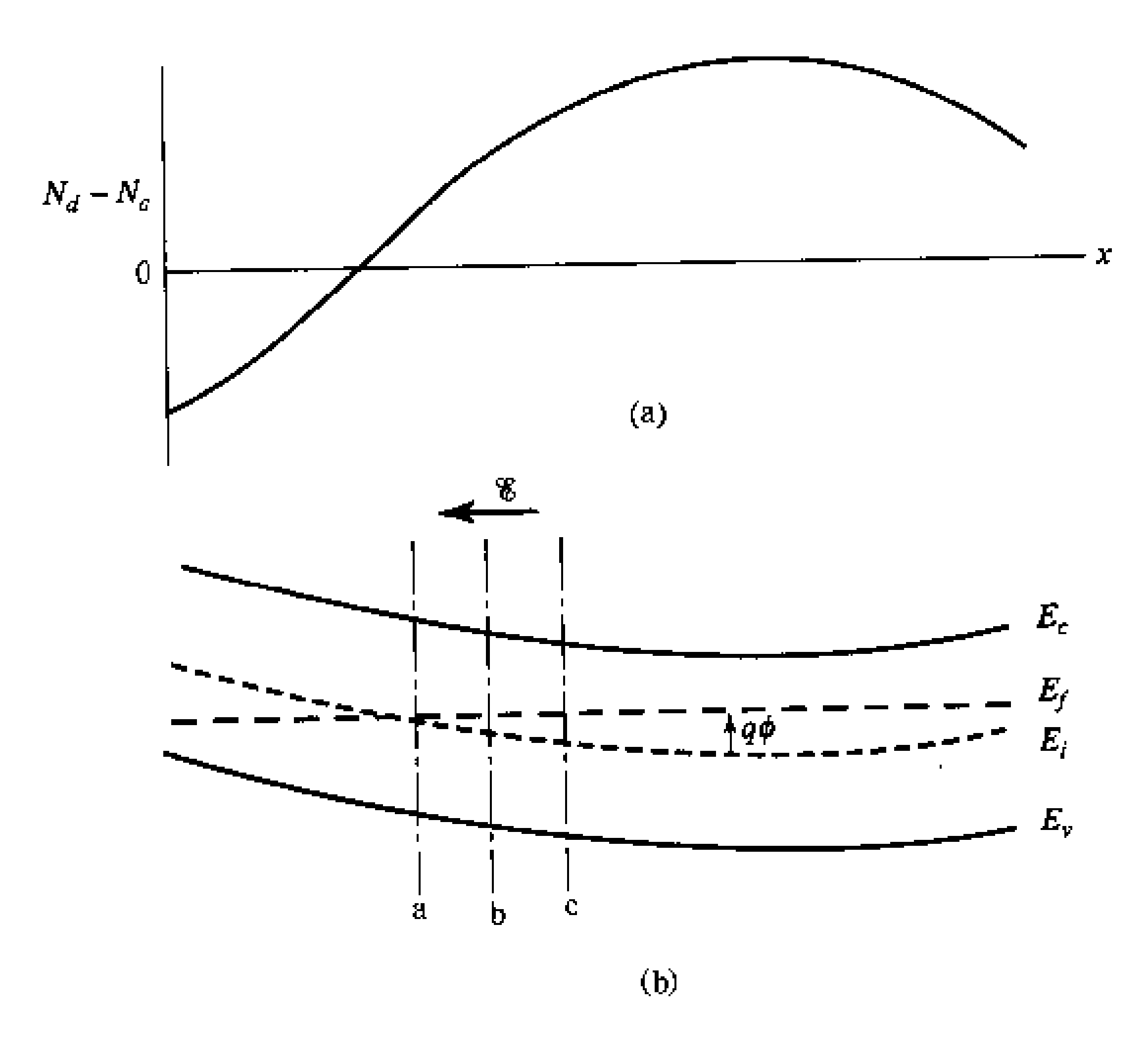


图4.1 (a)任意杂质分布半导体中净掺杂浓度与位置的关系。(b)对应的能带图,图中标出了电势 ϕ。文中将分析指定位置a、b和c处的情况

然而,由于杂质浓度和载流子浓度随位置变化,导致不同位置处Fermi能级与价带边、导带边的距离不同。图4.1(b)是与图4.1(a)的掺杂分布相应的能带图。Fermi能级与带边的距离在载流子浓度高的区域比载流子浓度低的区域小,本征Fermi能级 E_i 与Fermi能级 E_f 的交点,是净掺杂浓度 $N_d - N_a$ 等于零的位置。

电势 与上文讨论的载流子模型一样,电场的分布也可由能带图直接得出。图4.1(b)是电子的能带图,图上表示的电子能量从Fermi能级处算起。导带边和Fermi能级的距离反映了电子的势能,而高于导带边的能量则反映了电子的动能。在任何情况下,电势 ϕ 和电势能通过电荷 $-q$ 相联系,所以

$$\phi_c = -\frac{1}{q}(E_c - E_f) = \frac{1}{q}(E_f - E_c) \tag{4.1.1}$$

式中下标 c 表示相对于导带的能量。然而,势能的参考点是任意的,我们常用 E_i 取代 E_c 作为势能参考点,这时,电势 ϕ 不用下标。根据4.1(b),上式可写成

$$\phi = -\frac{1}{q}(E_i - E_f) = \frac{1}{q}(E_f - E_i) \tag{4.1.2}$$

根据这个定义,n型半导体的电势为正($E_f > E_i$),p型半导体的电势为负($E_f < E_i$)。

电场 因为电场等于电势的空间梯度的负值,从等式(4.1.2)中可得到电场 $\mathscr{E}_x$ 为

$$\mathscr{E}_x = -\frac{d\phi}{dx} = \frac{1}{q}\frac{dE_i}{dx} \tag{4.1.3}$$

所以,带边(和本征Fermi能级)的空间变化表明半导体中存在非零电场。在图4.1(b)的b

点,$\mathrm{d}\mathscr{E}/\mathrm{d}x$ 为负,电场指向左方,带负电荷的电子受到向右的作用力,因此,电子受到电场力作用,从高掺杂区 c 向低掺杂区 a 运动。p 型半导体的情况与此类似,只需改变符号和正负号。

现在我们来分析电场和缓变杂质分布之间的关系。一旦系统达到热平衡,半导体中的任意点不存在电流。此外,热平衡要求每个过程和它的反过程相平衡,所以这时电子电流和空穴电流都必须等于零。在 1.2 节中得到总的电子电流表达式为

$$J_n = q\mu_n n\mathscr{E}_x + qD_n\frac{\mathrm{d}n}{\mathrm{d}x} \tag{4.1.4}$$

这个表达式适用于 n 型半导体(电子是多数载流子)和 p 型半导体(电子是少数载流子)。

等式(4.1.4)的第一项代表漂移电流,第二项代表扩散电流。当总电子电流等于零时,这两部分电流相平衡,实际上没有电流流动,即每一点的漂移趋势和扩散趋势相互抵消。因为 $J_n=0$,可以根据电子浓度及其梯度解出

$$\mathscr{E}_x = -\frac{D_n}{\mu_n}\frac{1}{n}\frac{\mathrm{d}n}{\mathrm{d}x} = -\frac{kT}{q}\frac{1}{n}\frac{\mathrm{d}n}{\mathrm{d}x} \tag{4.1.5}$$

这里利用了等式(1.2.20)中定义的 Einstein 关系。同样,电场可以用空穴浓度表示,或者直接通过空穴电流的表达式(等式(1.2.22)),或者在等式(4.1.5)中利用质量作用定律(等式(1.1.13)),得到

$$\mathscr{E}_x = \frac{kT}{q}\frac{1}{p}\frac{\mathrm{d}p}{\mathrm{d}x} \tag{4.1.6}$$

等式(4.1.5)和等式(4.1.6)表明,如果已知可动载流子浓度及其梯度,就可以求出半导体中的电场。

在分析这个问题时,如果考虑到电子浓度与 Fermi 能级相对于带边距离(或者等价表示为本征 Fermi 能级)的关系,可以加深对物理概念的理解。考虑图 4.2(a)中位于 x_2 的电子,假定其能量为 E。这个能量中 $E-E_c$ 部分表示动能;其余部分代表势能。该电子可以在 x_1 与 x_4 之间的区域内自由运动,因为它的能量高于这一区域的势能。电子需要更高的能量才能抵达 x_1 左方或者 x_4 右方的区域,所以,经典理论认为电子不能抵达这些区域,存在阻止电子运动的势垒。

将材料中任意两点的载流子数目和能带结构相联系起来,因为在 x_3 处导带边与 Fermi 能级的距离比 x_2 处大,所以 x_2 处的电子浓度比 x_3 处低。将载流子浓度和电势 ϕ 用等式(4.1.5)联系起来。考虑到 $\mathscr{E}_x=-\mathrm{d}\phi/\mathrm{d}x$,半导体中任意一点有

$$\mathrm{d}\phi = \frac{kT}{q}\frac{\mathrm{d}n}{n} \tag{4.1.7}$$

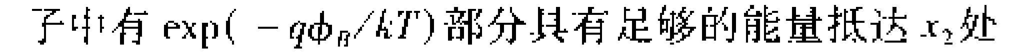
将这个等式在任意两点之间积分,比如从 x_2 到 x_3 得到

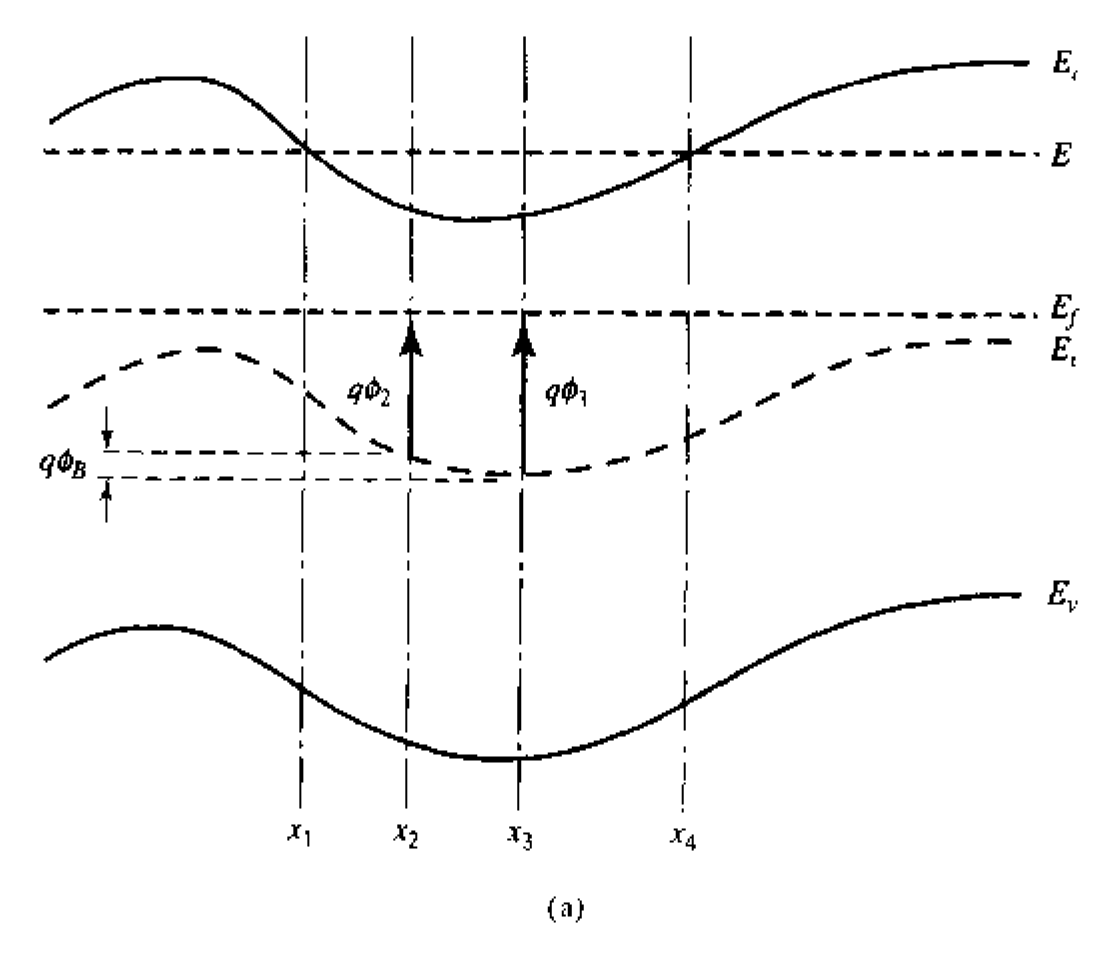
$$\phi_3 - \phi_2 = \frac{kT}{q}\ln\frac{n_3}{n_2} \tag{4.1.8}$$

将等式(4.1.8)改写成指数形式

$$\frac{n_3}{n_2} = \exp\left[\frac{q}{kT}(\phi_3 - \phi_2)\right] \tag{4.1.9}$$

载流子浓度之比取决于两点之间的势差 $\phi_B=\phi_3-\phi_2$。从物理上看可以理解为,位于 x_3 处的电


子中有 $\exp(-q\phi_B/kT)$ 部分具有足够的能量抵达 x_2 处



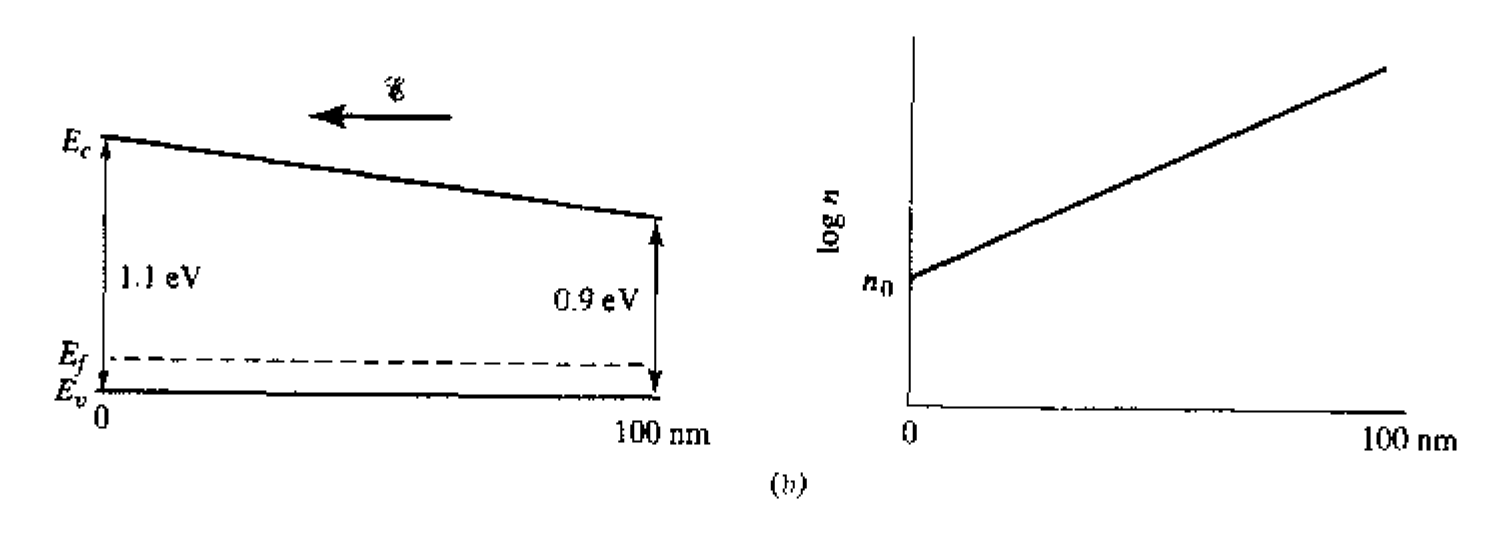


图 4.2　(a)任意掺杂半导体的能带图，具有能量 E 的电子被限制在 $E>E_i$ 的 x_1 到 x_4 之间的区域内运动；(b)带隙 E_g 随位置变化的 p 型材料的能带图及其相应的电子浓度分布

Poisson 方程　载流子浓度随位置变化时，常采用等式(4.1.9)。大部分的情况下，我们将 Poisson 方程[2]写成如下形式

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon_s} = -\frac{q}{\epsilon_s}(p - n + N_d - N_a) \tag{4.1.10}$$

式中 ρ 为空间电荷密度，假定杂质原子完全离化。在等式(1.1.26)中采用电势的定义(等式(4.1.2))，可将载流子浓度 n 和电势 ϕ 联系起来

$$n = n_i \exp\left(\frac{q\phi}{kT}\right) \tag{4.1.11}$$

这样，Poisson 方程就可以写成如下形式

---

2. 因为对半导体器件的许多分析与器件中不同区域间的载流子和电势的空间变化有关，正如第 3 章，经常要求解 Poisson 方程。再考虑到为了便于数学处理而采用的各种近似，解 Poisson 方程是半导体器件分析中最有用的方法之一。

$$\frac{d^2\phi}{dx^2}=\frac{q}{\epsilon_s}\left(2n_i\sinh\frac{q\phi}{kT}+N_a-N_d\right) \tag{4.1.12}$$

等式(4.1.12)是任意掺杂半导体电势分布的微分方程。不幸的是,通常情况下这个方程无法直接求解,必须采用近似才能得到特殊情况下的解析解。数值计算是另一种可以采用的求解途径。

为了进一步的讨论,考虑两种特殊情况:一种情况是掺杂浓度随位置缓变,比如扩散形成的n型区的施主分布;另一种情况与之相反,掺杂分布突变,例如p型半导体和n型半导体之间的结。

准中性 考虑掺杂浓度缓变的情形。在几百纳米尺度下分析掺杂浓度从$10^{18}cm^{-3}$变化到$10^{16}cm^{-3}$的n型硅,即典型的施主扩散的情况。这种施主浓度的变化对应的电势改变约为0.1V,电场(等式(4.1.5))约为$10^4V\ cm^{-1}$量级或更小。考虑如下特例,在0.5μm的范围内电场从0变化到$10^4V\ cm^{-1}$,平均电场梯度等于$2\times10^8V\ cm^{-2}$。把这个值代入Poisson方程(等式(4.1.10)),忽略少数载流子浓度p,我们发现n和N_d之差小于$10^{15}cm^{-3}$。因为这个数值在我们讨论的大部分区域都远小于施主浓度,所以将n近似为N_d进行分析是合理的。本质上,这种近似意味着多数载流子分布和施主杂质分布相差不大,这样半导体区域几乎是中性的,或者称为准中性。准中性近似在分析缓变掺杂浓度的情况更有效。在准中性近似的条件下,采用等式(4.1.5),n型半导体中的电场可以直接从施主浓度得到

$$\mathscr{E}_x=-\frac{kT}{q}\frac{1}{N_d}\frac{dN_d}{dx} \tag{4.1.13}$$

同样,中性近似下的p型半导体中的电场为

$$\mathscr{E}_x=\frac{kT}{q}\frac{1}{N_a}\frac{dN_a}{dx} \tag{4.1.14}$$

因为等式(4.1.13)和(4.1.14)成立的前提是准中性近似,所以当掺杂浓度梯度较大时不再适用

在采用准中性近似时,经常需要考虑指数掺杂分布的情形

$$N_d=N_0\exp\left(\frac{-x}{\lambda}\right) \tag{4.1.15}$$

式中λ是特征长度,反映了施主浓度随距离半导体表面$x=0$处的距离下降的快慢。为了简化数学计算,典型的Gauss分布或余误差分布(2.5节)经常被近似为指数分布,因为指数分布函数和它的导数——电场之间的关系比较简单,电场在整个指数掺杂区为常数$kT/q\lambda$。第6章中我们将看到,电场近似为常数简化了许多的器件分析。然而,指数近似可能会掩盖真实的扩散分布的某些重要细节。

例题 准中性近似

研究不均匀掺杂n型硅的准中性近似假设。已知掺杂浓度$N_d(x)$在$\lambda=1\mu m$的范围内从$10^{16}cm^{-3}$变化到$10^{18}cm^{-3}$,考虑这个区域的空间电荷分布。假设掺杂分布是

$$N_d(x)=10^{16}\times\exp\left\{\ln(100)\left[\frac{x}{\lambda}-\frac{1}{2\pi}\sin\left(\frac{2\pi}{\lambda}x\right)\right]\right\}\qquad\left(0<\frac{x}{\lambda}<1\right)$$

(这个数学表达式连续可微)[1]

计算 $0 < x/\lambda < 1$ 范围内的电场和电荷分布。

解：施主浓度随 x 的增加而增加，那么自由电子的浓度梯度为正，因此，电子趋于沿负 x 方向扩散。热平衡时，这种扩散趋势必须与电子沿正 x 方向的漂移趋势达到平衡，这就需要一个指向负 x 方向的内建电场。与这个电场相联系的是空间电荷，所以缓变掺杂区域并不是真正的电中性。本例题将计算电荷密度，以定量研究与电中性假定的偏离。

利用等式(4.1.8)，计算出跨越整个区域的电势差 $\Delta\phi$ 为

$$\Delta\phi = |\mathscr{E}_{avg}|\lambda = \frac{kT}{q}\ln(100) = 0.12\ \mathrm{V}$$

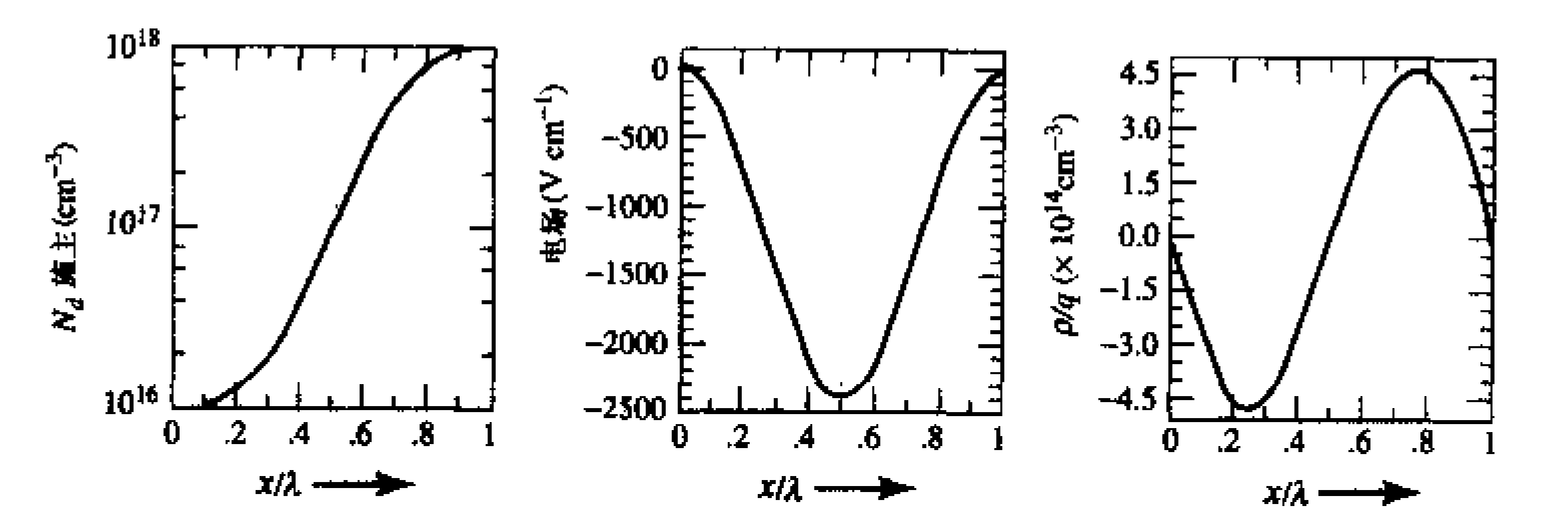
式中 $\mathscr{E}_{avg} = -1.19\times10^3\ \mathrm{V\ cm^{-1}}$ 是该区域中的平均电场。从等式(4.1.13)可以看出，电场是 x 的函数

$$\mathscr{E} = -\frac{kT}{q}\frac{1}{N_d}\frac{\mathrm{d}N_d}{\mathrm{d}x} = \mathscr{E}_{avg}\left(1-\cos\frac{2\pi}{\lambda}x\right)$$

为了求出电荷密度，用 Poisson 方程 $\rho(x) = \epsilon_s \mathrm{d}\mathscr{E}/dx$。用电荷归一化得到

$$\frac{\rho(x)}{q} = \frac{\epsilon_s\mathscr{E}_{avg}}{q\lambda}\left[2\pi\sin\left(\frac{2\pi}{\lambda}x\right)\right]$$

在 $x/\lambda = 1/4$ 和 $x/\lambda = 3/4$ 处归一化电荷密度($|\rho/q|$)最大，等于 $4.8\times10^{14}cm^{-3}$，比最小掺杂浓度($N_d = 10^{16}cm^{-3}$)略小一些。因此，我们可以得出以下结论，在这个掺杂浓度缓变的区域采用准中性近似是合理的。下面附图描述了掺杂缓变区域的杂质浓度、电场和电荷密度分布。



异质材料 迄今为止，一直假设半导体材料的性质不随着空间位置变化。然而，在某些情况下，为了改善器件性能，会刻意改变半导体的组分(和带隙)。甚至整个系统的掺杂浓度不变时，带隙变化所产生的电场也会加速或阻碍自由载流子的运动。

为了便于说明，考虑一块尺度在 100nm，带隙从 1.1eV 线性变化到 0.9eV 的 p 型的半导体，如图4.2(b)所示。空穴和电子浓度可写成

$$p = N_v \exp\frac{-(E_f - E_v)}{kT} \tag{4.1.16}$$

和

$$n = N_c \exp\frac{-(E_c - E_f)}{kT} \tag{4.1.17}$$

式中 N_i 和 N_c(等式(1.1.23)和(1.1.24))是位置的函数。然而,如果各材料的能带结构变化不太大,N_i 和 N_c 随组分或者位置也不会有大的变化

对于 p 型材料,$p \approx N_a$。因为 $E_f - E_i \approx (kT/q)\ln(N_a/N_i)$,价带边近似平行于统一的 Fermi 能级。然而,由于带隙随位置改变,导带边以及相应的电子浓度也发生变化。在例题中采用等式(4.1.16)和(4.1.17)得到的带隙变化是 0.2eV,我们发现

$$n \approx n_0 \exp\left(\frac{0.2\ \text{eV}}{0.026\ \text{eV}}\frac{x}{100\ \text{nm}}\right) = n_0 \exp\left(\frac{x}{13\ \text{nm}}\right) \tag{4.1.18}$$

式中 n_0 是 $E_g = 1.1\text{eV}$ 处的电子浓度。从式中可看出,带隙小的地方电子浓度高。

与掺杂梯度类似,带隙的梯度变化将引入电场以平衡电子从高浓度的区域向低浓度的区域扩散的趋势。电场的方向是将电子推回到高浓度的区域,所以必然是负的。这一电场对注入到材料中的过剩电子具有非常重要的作用。因为这个电场是负的,从左侧注入的过剩电子将被加速。电子穿越组分缓变区域的速度比穿越恒定组分的区域快,这种加速电场对晶体管速度的影响将在 7.6 节中讨论。

4.2 pn 结

上节只讨论了一种导电类型,并且掺杂浓度随位置缓变的材料。在施主浓度或组分变化的准中性区域由这些约束条件可得到电场的解。现在,考虑另一个极端:半导体的掺杂浓度或组分突变。这种情况下,半导体局部区域将严重地偏离电中性条件。考虑 p 型和 n 型半导体形成的结,我们发现过渡区可动载流子被耗尽。这一耗尽近似是与准中性近似相反的另一个极端。在分析器件结构时,通常假定某些区域为准中性区,其余的区域可动载流子完全耗尽。尽管这只是一种理想情况,但这种简化对大部分计算已经足够了。更精确的分析需要采用数值计算。

为了建立 pn 结模型,首先考虑空间分离的同种材料:n 型和 p 型半导体,如图 4.3(a)所示。当它们密切接触时,如图 4.3(b)所示,两个区域中电子浓度的巨大差异将导致电子从 n 型半导体流向 p 型半导体,空穴从 p 型区流向 n 型区。当这些可动载流子运动到另一种掺杂类型的区域时,靠近结的区域上将留下未被补偿的杂质离子,形成了电场。电力线从结的 n 型一侧指向 p 型一侧,如图 4.3(c)所示,在两个区域之间产生势垒。达到平衡时,电场的大小恰好使得电子从 n 型区向 p 型区扩散的趋势与内建电场下电子向相反方向漂移的趋势抵消。

势垒 利用两个区域未接触时的(见图 4.3(a))Fermi 能级差可确定与内建电场相关的势垒的大小,正如第 3 章中对金属-半导体系统的处理。当密切接触的半导体达到平衡时,整个系统的 Fermi 能级必须一致。因此,两个区域间形成的能量势垒必定等于它们分离时的 Fermi 能级之差。这个值等于分立的半导体的功函数之差,因为半导体的功函数的定义为

$$q\Phi_s = q\text{X} + (E_c - E_f) \tag{4.2.1}$$

式中 $q\text{X}$ 是电子亲和能。

在远离结的区域,载流子的浓度与孤立的半导体相同。n 型半导体电子浓度等于施主浓度 N_d,空穴浓度等于 n_i^2/N_d。同样,远离结的 p 型半导体空穴浓度等于 N_a,电子浓度是 n_i^2/N_a。因为远离结的区域的载流子浓度已知,可以根据等式(4.1.11)解出由等式(4.1.2)定义的电势。在结附近的区域,我们却不容易确定自由载流子浓度。然而,图 4.3(c)表明这一区域的 $|\phi|$ 是很小的。由等式(4.1.11)知当 ϕ 变小时,载流子浓度迅速减小,因此,过渡区的载流子浓度远小于中性区。因此,过渡区经常被称为耗尽区,该区域的空间电荷基本上全部由杂质离子组成。

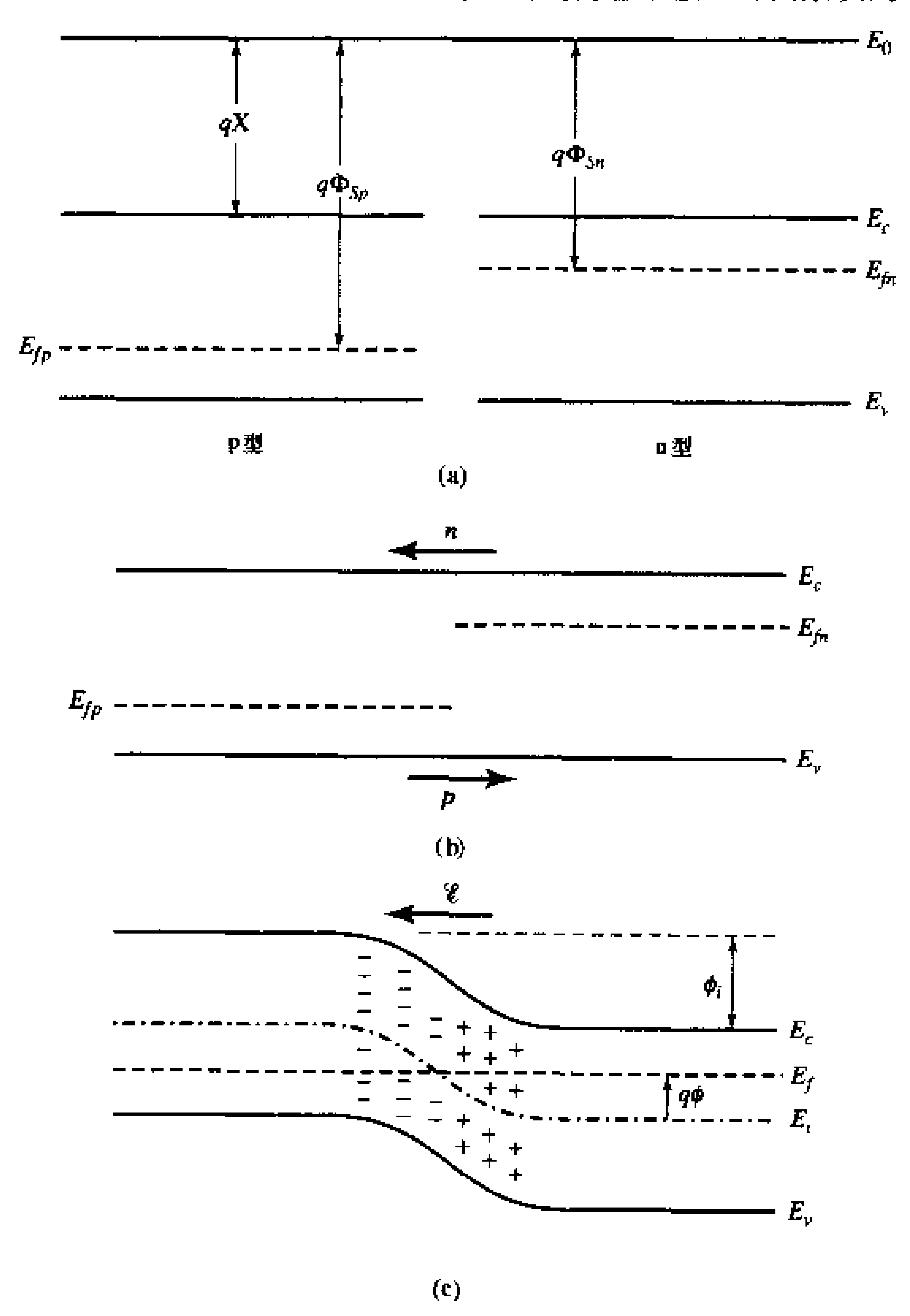


图4.3 (a)空间分离的 n 型和 p 型半导体,没有达到热平衡;(b)两种材料密切接触,使得空穴从 p 型区向 n 型区扩散、电子从 n 型区向 p 型区扩散;(c)自由载流子的迁移形成未被补偿的杂质离子,产生的电场阻止(和平衡)空穴与电子的扩散趋势

耗尽近似假设半导体能够明确地划分成中性区和可动载流子完全耗尽区。这两个区域在耗尽区或空间电荷区的边界上衔接,假设在那里多数载流子浓度从施主浓度突变到零。这种耗尽近似简化了 Poisson 方程(等式(4.1.10))的求解。

因为假定在耗尽区载流子浓度远小于离化的杂质浓度,对电势二次微分正比于耗尽区的净掺杂浓度

$$\frac{d^2\phi}{dx^2} = \frac{-q}{\epsilon_s}(N_d - N_a) \qquad (4.2.2)$$

通常,N_d和N_a可能是位置的函数,这样就难以直接求出方程(4.2.2)的解析解。不过,考虑一些理想化的掺杂分布,就能求解出方程(4.2.2),并且得到关于实际pn结的一些物理概念。我们先考虑突变结。突变结是这样一种结构:在结的一侧是均匀掺杂的n型区,在结处突变到均匀掺杂的p型区。正如第2章讨论的,突变结可以通过在p型衬底上外延均匀掺杂的n型区来形成。

然后考虑线性缓变结。这种结构的特点是,杂质浓度从n型区到p型区线性变化。在限定的区域内,某些通过扩散形成的结可以近似为线性缓变结。

最后考虑异质结结构,它由两种不同带隙的半导体材料构成n型区和p型区。在高速晶体管和光学器件中,异质结非常重要。它们可由两种不同的Ⅳ族材料组成,或者由Ⅲ族和Ⅴ族元素形成的不同的化合物半导体组成。正如2.8节讨论的,异质结通常是在一种半导体材料上外延晶格常数相近(保证外延层无应力)的另一种半导体材料构成。

4.2.1 突变结

近似分析 考虑图4.4所示的突变结,掺杂浓度在$x=0$处从N_a突变到N_d,我们来解方程(4.2.2)。采用耗尽近似,假设从$-x_p$到x_n之间的区域可动载流子全部耗尽,如图4.4(b)所示。可动的多数载流子浓度在耗尽区边缘与掺杂浓度相等。所以除了耗尽区,电荷密度都等于零,耗尽区电荷密度则等于离化的杂质浓度,如图4.4(c)所示。在n型半导体($x>0$)中,等式(4.2.2)变为

$$\frac{d^2\phi}{dx^2} = -\frac{d\mathscr{E}}{dx} = -\frac{qN_d}{\epsilon_s} \qquad (4.2.3)$$

这个式子很容易求出从n型耗尽区任意一点到耗尽区边缘x_n的积分。在x_n处半导体变成电中性,电场消失。完成这一积分,得到的电场是

$$\mathscr{E}(x) = -\frac{qN_d}{\epsilon_s}(x_n - x) \qquad 0 < x < x_n \qquad (4.2.4)$$

整个耗尽区电场是负值,随x线性变化,在$x=0$处达到最大值,如图4.4(d)所示。从物理上看,电场的方向指向左侧是合理的,因为它作用在带负电荷的电子上的力必须和电子从电中性n型半导体向左侧扩散的趋势相抵消。用类似的方法可以得到p型区的电场

$$\mathscr{E}(x) = -\frac{q}{\epsilon_s}N_a(x + x_p) \qquad -x_p < x < 0 \qquad (4.2.5)$$

p型区的电场也是负的,同样是为了阻止带正电的空穴向右侧扩散

在$x=0$处电场必须连续,所以

$$N_a x_p = N_d x_n \qquad (4.2.6)$$

因此,结任意一侧的耗尽区宽度与掺杂浓度成反比,即掺杂浓度越高,空间电荷区越窄。考虑某一侧的掺杂浓度远高于另一侧的不对称结,耗尽区主要延伸在轻掺杂区,而重掺杂一侧的耗尽区

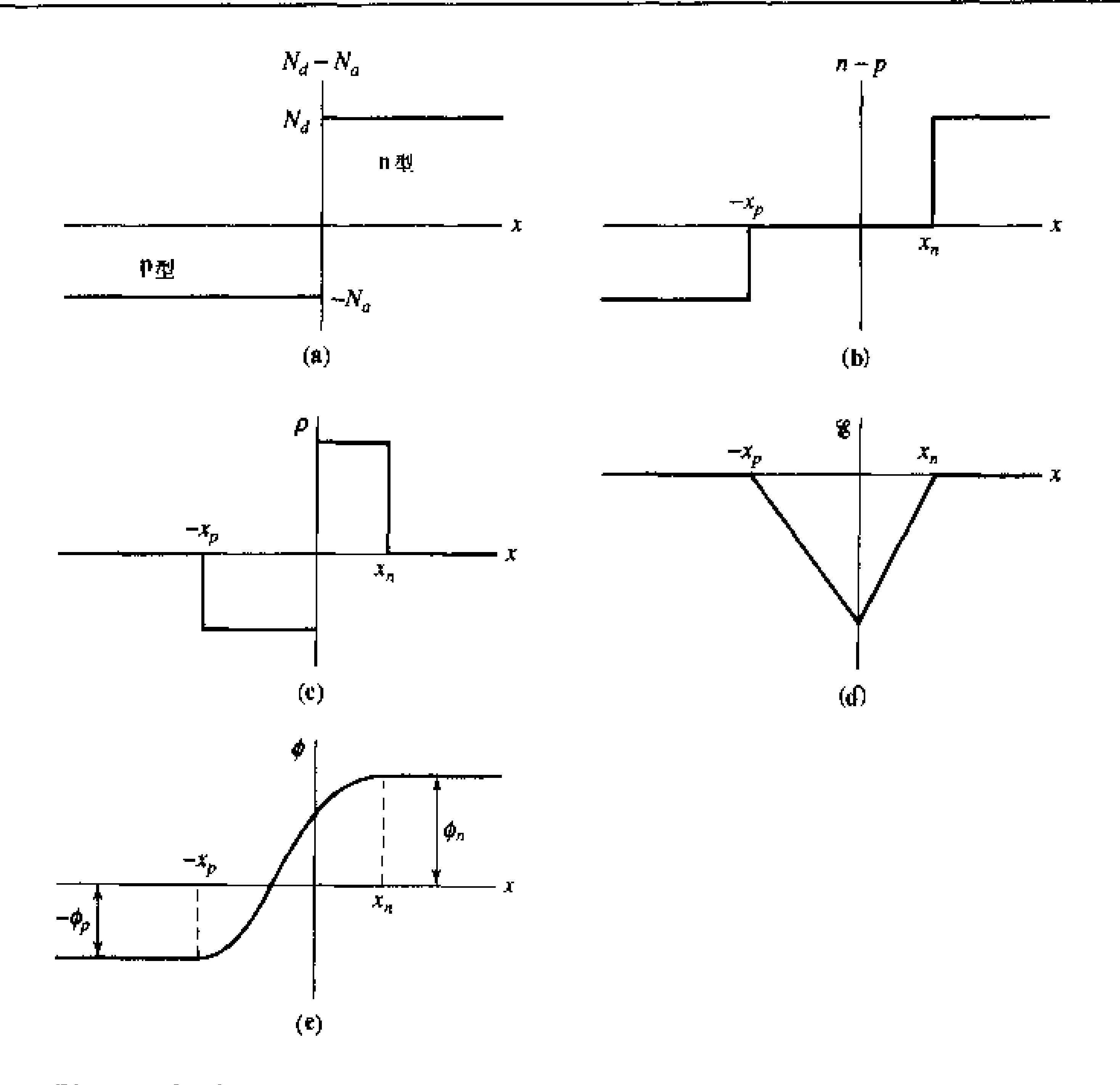


图4.4 采用耗尽近似分析得到的突变结性质：(a)净掺杂浓度分布；(b)载流子浓度；(c)Poisson 方程采用的空间电荷分布；(d)对 Poisson 方程一次积分得到的电场分布；(e)两次积分得到的电势分布

宽度可以被忽略。这种被称做单边突变结的电荷、电场、电势分布和理想的 Schottky 势垒相同，相应的图4.4可简化成图3.4。单边突变平面结的性质经常被用到，其列线图如表4.1所示。

对电场的表达式再次积分，得到结上电势分布。对 n 型半导体

$$\phi(x) = \phi_n - \frac{qN_d}{2\epsilon_s}(x_n - x)^2 \qquad 0 < x < x_n \tag{4.2.7}$$

如图4.4(e)所示，ϕ_n是从等式(4.1.11)求出的耗尽区的中性边界的电势

$$\phi_n = \frac{kT}{q}\ln\frac{N_d}{n_i} \tag{4.2.8}$$

同样，对 p 型半导体

$$\phi(x) = \phi_p + \frac{qN_a}{2\epsilon_s}(x + x_p)^2 \qquad -x_p < x < 0 \tag{4.2.9a}$$

$$\phi_p = \frac{-kT}{q}\ln\frac{N_a}{n_i} \tag{4.2.9b}$$

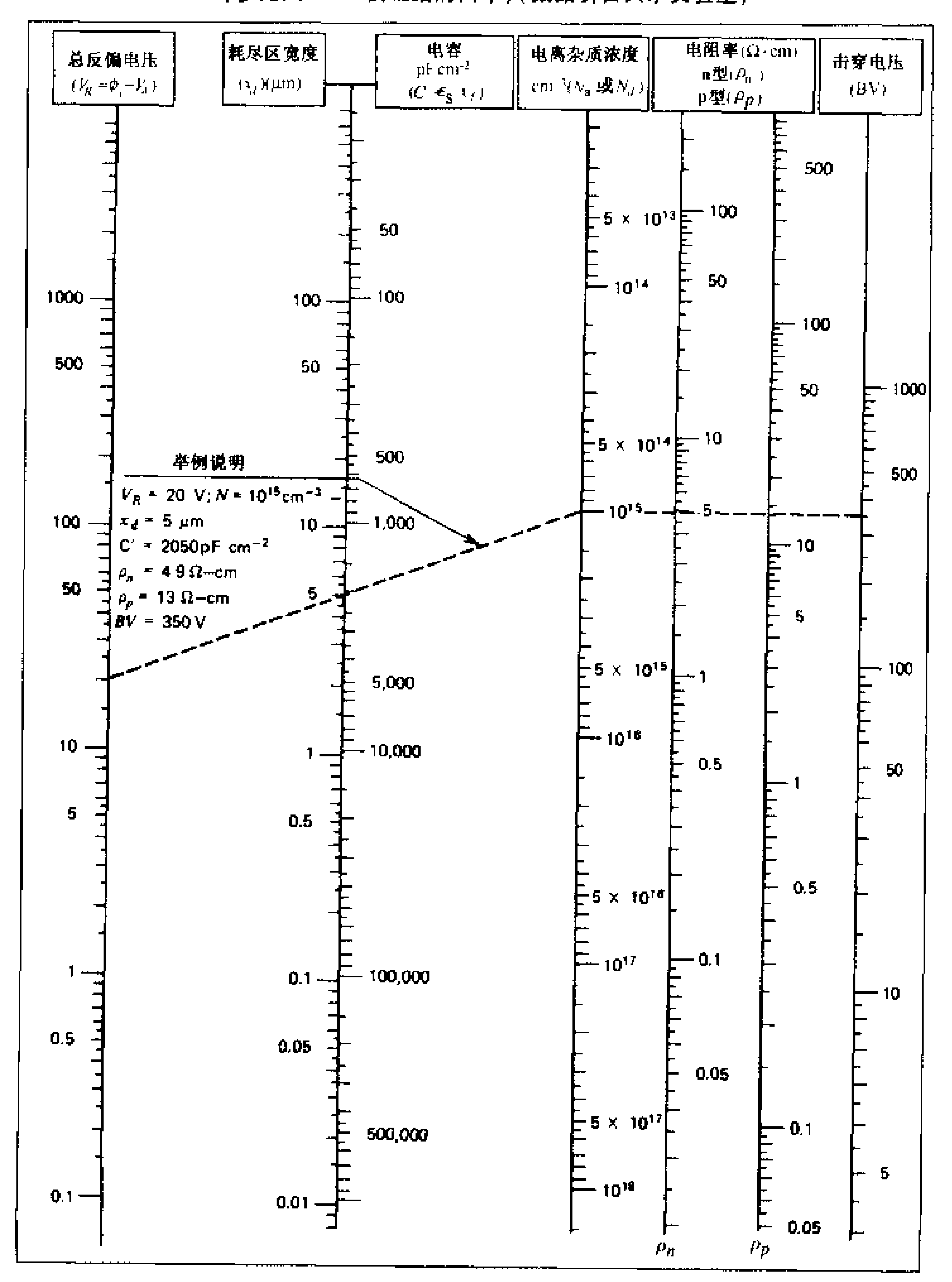
式中 $\phi_p<0$,是 p 型材料的耗尽区的中性边界处的电势。

从中性 p 型区到中性 n 型区的总的电势差 ϕ_t 等于 $\phi_n-\phi_p(\phi_p<0)$,取决于两个区域的掺杂浓度

$$\phi_t=\phi_n-\phi_p=\frac{kT}{q}\ln\frac{N_d}{n_i}+\frac{kT}{q}\ln\frac{N_a}{n_i}=\frac{kT}{q}\ln\frac{N_dN_a}{n_i^2} \tag{4.2.10}$$

表 4.1 均匀掺杂的硅单边突变结各种参数列线图(300K)

(参见图 4.15 校正结的曲率)(数据引自贝尔实验室)



注意，内建电势 ϕ_i 是正值（也就是说，n 型一侧的电势比 p 型一侧的要高），这样正好可以达到漂移运动和扩散运动的平衡。电势变化的主要部分发生在低掺杂区，而且低掺杂区的耗尽区也比较宽。还需注意的是，除非结是对称的（例如，$N_a=N_d$），结平面（$x=0$）处的电势并不为零。

例题　求 pn 结中的内建电势

已知电阻率为 4Ω · cm 的轻掺杂的 n 型硅样品，与掺杂浓度 N_a 是其1000 倍的 p 型硅形成 pn 结，求结上的内建电势 ϕ_i。

解：等式(4.2.10)将 ϕ_i 表示成结两侧的掺杂浓度 N_a 和 N_d 的函数。从图 1.15 得到电阻率为 4Ω · cm 的 n 型硅样品对应的 $N_d=10^{15}\text{cm}^{-3}$。因为 p 型区的掺杂浓度是 n 型区的 1000 倍，所以 $N_a=10^{18}\text{cm}^{-3}$。

$$\phi_i=\frac{kT}{q}\ln\left(\frac{N_aN_d}{n_i^2}\right)=\frac{kT}{q}\ln\left(\frac{10^{18}\times10^{15}}{(1.45\times10^{10})^2}\right)$$
$$\phi_i=0.753\text{ V}$$

这一电压约等于 E_g/q 的 2/3，是多数 IC 中 pn 结 ϕ_i 的典型值。

掺杂浓度非常高时，基于 Maxwell-Boltzmann 统计分布（例如，等式(4.2.8)）的等式(4.2.10)不再有效。当掺杂浓度接近有效态密度 N_c 或 N_v（10^{19}cm^{-3}）时，必须考虑 Fermi-Dirac 统计。然而，在计算 pn 结上的电势时，不需要仔细考虑 Fermi-Dirac 统计，因为高掺杂浓度时，Fermi 能级非常接近带边，重掺杂硅一侧的电势大约等于 $E_g/2q$，或 0.56V。所以，由重掺杂的 p 型硅（通常表示成 p^+ 硅）和轻掺杂的 n 型硅组成的 pn 结的内建电势是

$$|\phi_i|=0.56+\frac{kT}{q}\ln\left(\frac{N_d}{n_i}\right)$$

对 n^+ 硅和轻掺杂的 p 型硅形成的 pn 结可得到相似的结论，只需用 N_a 取代 N_d 即可。

对任意掺杂浓度的突变 pn 结，从等式(4.2.6)到(4.2.10)得到的总耗尽区宽度为

$$x_n+x_p=\left[2\frac{\epsilon_s}{q}\phi_i\left(\frac{1}{N_a}+\frac{1}{N_d}\right)\right]^{1/2} \tag{4.2.11}$$

从等式(4.2.11)可看出，耗尽区宽度主要取决于轻掺杂区，并近似与轻掺杂浓度的平方根成反比。

例题　反偏突变结

考虑均匀掺杂的 pn 结，p 型区的掺杂浓度为 N_a，n 型区的掺杂浓度为 N_d。

如果外加反偏电压 $V_a=-5\text{V}$，推导降落在 n 型区的反偏电压占结上总反偏电压的百分比 P_n 的表达式。对于 $N_a=10^{17}\text{cm}^{-3}$ 的 pn 结，计算以下三种情况的 P_n：(a) $N_d=10^{-1}N_a$，(b) $N_d=10^{-2}N_a$，(c) $N_d=10^{-3}N_a$。

利用等式(4.2.11)的在外加偏压时的修正形式，求出各种情况的耗尽层宽度。将计算结

果与查表4.1得出的耗尽层宽度 x_d 进行比较。

解：本题的电场分布如图4.4(d)所示。用 $\mathscr{E}_{max}$ 表示最大的电场值，降落在p型区的电压 V_p 为

$$V_p = \frac{1}{2}\mathscr{E}_{max} x_p$$

同样，降落在n型区的电压 V_n 为

$$V_n = \frac{1}{2}\mathscr{E}_{max} x_n$$

从这两个式子得到 $V_p/V_n = x_p/x_n$。结上总的反偏电压是 $V_R = V_p + V_n$。所以求出的百分比 P_n 为

$$P_n = \frac{V_n}{V_n + V_p} \times 100 = \frac{1}{1 + V_p/V_n} \times 100$$
$$= \frac{1}{1 + x_p/x_n} \times 100$$

将等式(4.2.6)代入，得到

$$P_n = \frac{1}{1 + N_d/N_a} \times 100$$

利用该式，求出这三种情况的百分比(P_n值)为

(a) $N_d = 10^{16}$　　$P_n = 91\%$

(b) $N_d = 10^{15}$　　$P_n = 99\%$

(c) $N_d = 10^{14}$　　$P_n = 99.9\%$

因此突变结两侧掺杂浓度相差一个数量级时，大于90%的反向偏压降落在轻掺杂一侧。当掺杂浓度比增加时，就更接近于单边结。

本题的第二部分考虑结上加5V反偏电压时的总耗尽层宽度。利用等式(4.2.11)的修正形式，用结上的总反偏电压 $V_R = \phi_i + |V_a|$(p型区相对于n型区)替换掉原来的内建电势 ϕ_i。根据等式(4.2.10)，对于情况(a)、(b)和(c)，得到 ϕ_i 为

(a)对于 $N_a = 10^{17}\text{cm}^{-3}$, $N_d = 10^{16}\text{cm}^{-3}$, $\phi_i = 0.753\text{V}$

(b)对于 $N_a = 10^{17}\text{cm}^{-3}$, $N_d = 10^{15}\text{cm}^{-3}$, $\phi_i = 0.694\text{V}$

(c)对于 $N_a = 10^{17}\text{cm}^{-3}$, $N_d = 10^{14}\text{cm}^{-3}$, $\phi_i = 0.634\text{V}$

因为 $V_a = -5\text{V}$，对于情形(a)、(b)和(c)分别得到 $V_R = 5.753\text{V}$、5.694V和5.634V。在修正后的等式(4.2.11)和表4.1中代入这些值，得到的 x_d 为

	等式(4.2.11)	表4.1
(a)	0.91μm	0.85μm
(b)	2.73μm	2.7μm
(c)	8.54μm	8.6μm

本例题的 pn 结在反偏时明显表现出单边结的特征，甚至当两边掺杂浓度比值较小时也是这样。即使在两边掺杂浓度只差一个数量级时，表 4.1 对于粗略的估算耗尽层宽度也是很有用的。

突变 pn 结的分析是采用耗尽近似来确定边界条件并求解 Poisson 方程的。在继续讨论之前，我们先简要讨论一下更精确的解，以及与耗尽近似的结果的比较。重新考虑 $x=0$ 处掺杂浓度突变形成的结，如图 4.5(a) 所示。不再假定 x_n 和 $-x_p$ 处由准中性区突变到完全耗尽区，而是认为这些边界上的过渡区域只是部分耗尽，如图 4.5(b) 所示。因为过渡区（见图 4.5(c)）的净电荷密度比掺杂浓度低，所以与耗尽近似相比，电场的变化要缓慢一些。图 4.5(d) 的实线代表精确解的电场，虚线则代表耗尽近似得到的电场。更精确的分析表明，电场向半导体内部进一步延伸，空间电荷区则更宽。

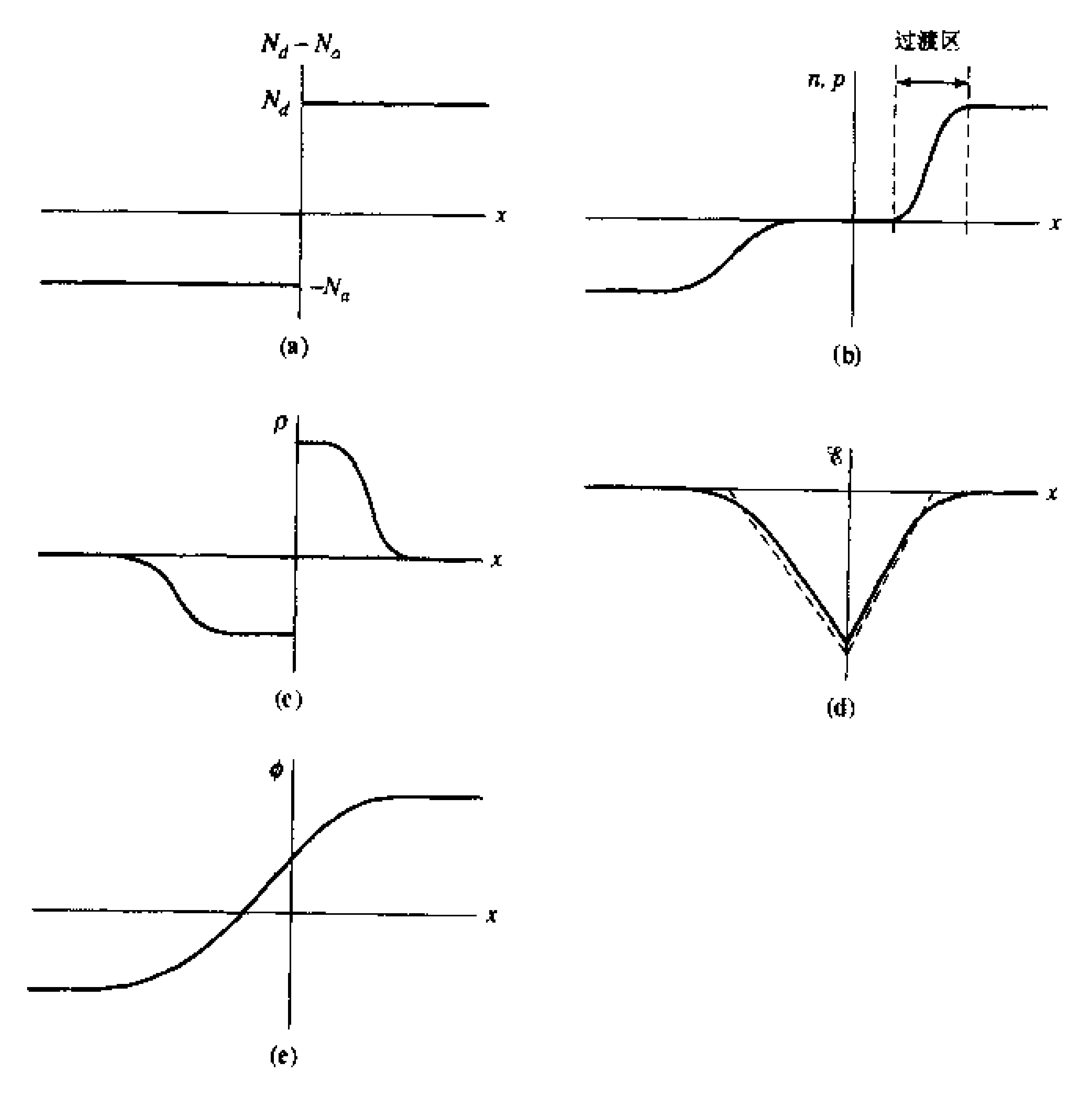


图 4.5 考虑中性区和耗尽区边界缓变时的突变结性质：(a) 净掺杂浓度分布；(b) 载流子浓度分布；(c) 空间电荷分布；(d) 电场分布；(e) 电势分布（与图 4.4 比较）

Debye 长度[†] 尽管不采用耗尽近似也能够得到突变结的精确解，但是一般不采用这种方法分析，而是在空间电荷区边界区域（x_n 和 $-x_p$ 附近）采用电势的近似解析表达式来研究耗尽近似的有效性。假定电势在 $x=x_n$ 附近与 ϕ_n 有较小的偏离，忽略少数载流子浓度 p，并在等式 (4.1.11) 中令 $\phi'=\phi_n-\phi$，将等式 (4.1.10) 改写为

$$\frac{d^2\phi'}{dx^2} = \frac{q}{\epsilon_s}(N_d - n) = \frac{q}{\epsilon_s}\left[N_d - n_i \exp\left(\frac{q(\phi_n - \phi')}{kT}\right)\right] \qquad (4.2.12)$$
$$= \frac{q}{\epsilon_s}N_d\left[1 - \exp\left(-\frac{q\phi'}{kT}\right)\right]$$

因为 ϕ' 很小,可以将等式(4.2.12)中的指数项台劳展开,保留前两项,那么等式就简化为

$$\frac{d^2\phi'}{dx^2} = \frac{q}{\epsilon_s}N_d\frac{q\phi'}{kT} = \frac{\phi'}{L_D^2} \qquad (4.2.13)$$

式中 L_D 是非本征 Debye 长度

$$L_D = \left[\frac{\epsilon_s kT}{q^2 N_d}\right]^{1/2} \qquad (4.2.14)$$

这是与电势的空间变化相关的特征长度[3]。在讨论 Schottky 欧姆接触时,我们介绍了类似的 Debye 长度公式,即等式(3.4.3)。等式(4.2.13)解的形式为 $\phi' = B\exp(x/L_D)$,其中 B 是积分常数。所以电势 ϕ' 在空间电荷区的边界随 x 指数变化,其特征长度等于非本征 Debye 长度。因为载流子浓度和电势呈指数变化关系,仅在几个 Debye 长度下,载流子浓度就能从掺杂浓度快速下降到零。因此,耗尽近似仅在空间电荷区边缘 x_n 和 $-x_p$ 附近几个非本征 Debye 长度区域内才需要修正。对于典型掺杂浓度为 10^{16}cm^{-3} 的对称结,L_D 等于 40nm,然而求解等式(4.2.6)和(4.2.11)得到的 x_n 近似为 210nm。因此,在这种情况下,耗尽近似是合理的,但显然只是一种近似。

4.2.2 线性缓变结

通常,用突变结表示扩散形成的 pn 结并不合适,并且对大多数实际的双扩散结来说,更不适用。双扩散结是相继扩散两种相反类型的掺杂原子构成的结。用 Poisson 方程(4.1.12)不能求出解析解,但是对掺杂分布进行简化,则可以得到近似的解析解。如果想得到更精确的解,就必须采用数值计算。

扩散形成的 pn 结的一种近似是线性缓变结,采用这种近似能够精确求解。

线性缓变结的净掺杂浓度从 p 型区到 n 型区线性变化,用常数 a 表征,即净掺杂浓度梯度(单位是 cm^{-4})。这样,整个空间电荷区(见图 4.6(a))的净掺杂浓度就可以写成

$$N_d - N_a = ax \qquad (4.2.15)$$

---

③ 更严格地讲,Debye 长度 L_D 描述了可动载流子重新分布对电场的屏蔽,它取决于区域中总的自由载流子浓度(空穴和电子).

$$L_D = \left[\frac{\epsilon_s kT}{q^2(n+p)}\right]^2$$

对于非本征 n 型半导体,少数载流子浓度可忽略,L_D 采用等式(4.2.14)的形式. 当载流子浓度下降时,Debye 长度增加。Debye长度的最大值就是由下式给出的本征Debye长度 L_{Di}

$$L_{Di} = \left[\frac{\epsilon_s kT}{2q^2 n_i}\right]^{1/2}$$

室温下,$L_{Di} = 24\mu\text{m}$。

采用耗尽近似，电场和电势很容易从 Poisson 方程解出。因为耗尽区的空间电荷随位置线性变化，所以在空间电荷区，电场是位置的二次函数，电势是位置的三次函数，如图 4.6 所示（详细分析见习题 4.6）。

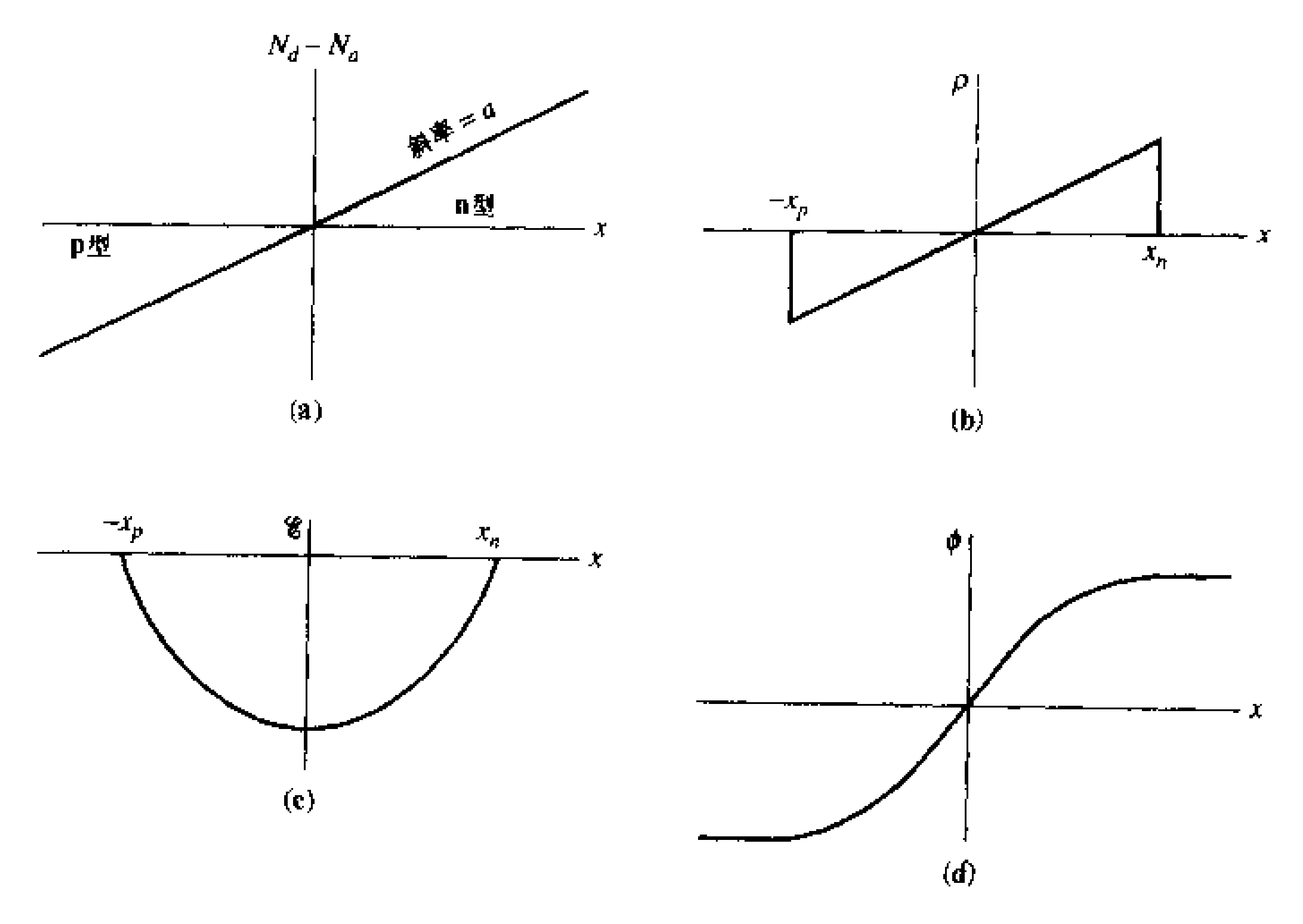


图 4.6　采用耗尽近似分析得到的线性缓变结的性质：(a) 净掺杂浓度分布：$N_d - N_a = ax$；(b) 空间电荷分布；(c) 电场分布；(d) 电势分布

尽管实际上不可能实现线性缓变结，但是在有限的电压范围内，很多实际情况可以近似成线性缓变结。如果突变结被加热，那么杂质原子将扩散通过结，结区变得不再陡峭，这样只要空间电荷区的宽度小于杂质原子的扩散长度，就可以被近似为线性缓变结。如图 4.7(a) 所示的 p 型衬底上扩散 n 型杂质形成的扩散结，在有限的距离内，有时也被近似成线性缓变结。另一方面，在均匀掺杂衬底上进行杂质扩散，如果扩散长度小于空间电荷区宽度，如图 4.7(b) 所示，则可近似成单边突变结，或者说是非理想扩散形成了"箱状"的分布。

指数掺杂　尽管更多的真实的结不能完全解析求解，但是也可以对它们做一些定性分析。在相当大的范围内可以将余误差分布或 Gauss 分布近似为指数分布。这样，在均匀掺杂的 p 型衬底上扩散 n 型杂质得到的净掺杂分布，如图 4.8(a) 所示，可以被写成

$$N_d - N_a = N_0 e^{-x/\lambda} - N_a \tag{4.2.16}$$

式中，λ 是与扩散相关的特征长度，N_a 是 p 型衬底的掺杂浓度。正如 4.1 节指出的，在指数掺杂的准中性 n 型区将存在电场，以平衡载流子向低浓度区扩散；而均匀掺杂的 p 型中性区没有电场。结的 n 型一侧耗尽区内没有可动载流子，因此这里的电场比准中性的 n 型区要高，图 4.8(b) 给出了整个结的电场分布。因为准中性区的电场是常数，而空间电荷可由电场的微分求出，所以那里的空间电荷为零，尽管在表面还存在一层很薄的面电荷，如图 4.8(c) 所示。其余的空间电荷将被限制在耗尽区内。

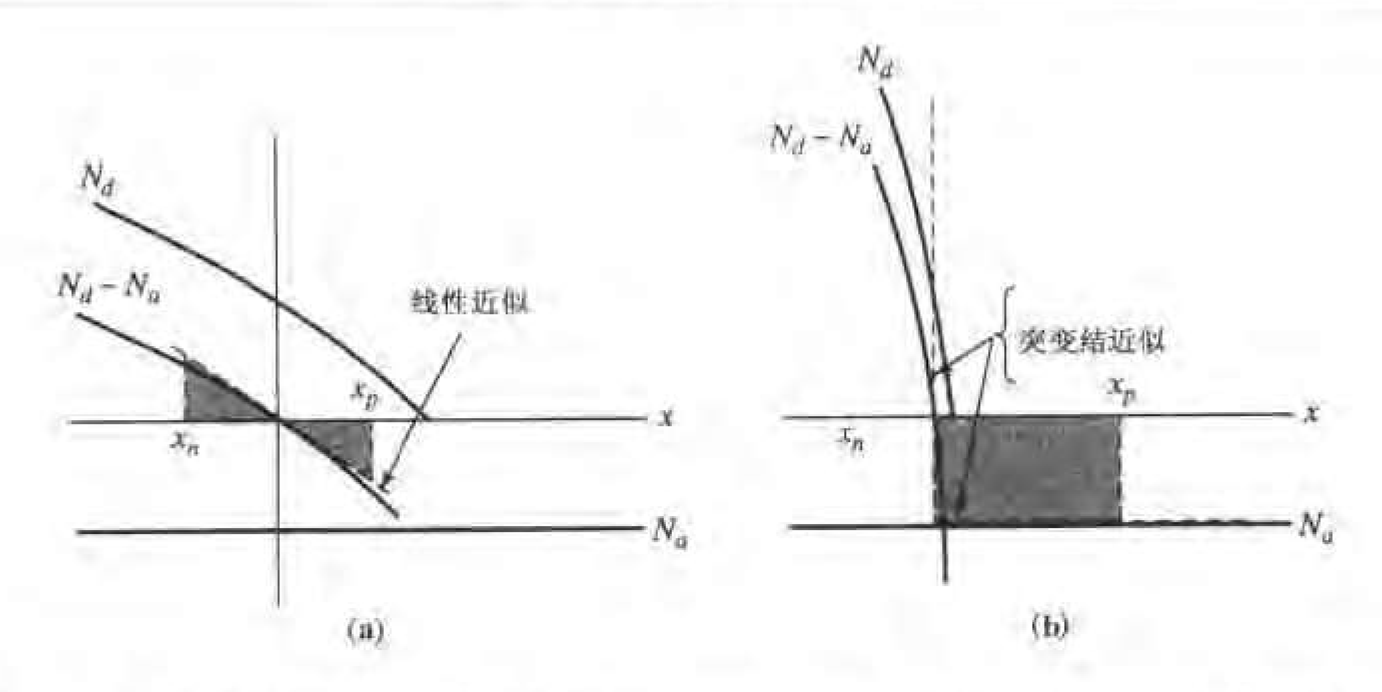


图 4.7　(a)如果扩散长度 $2\sqrt{Dt}$ 远大于空间电荷区宽度,扩散结可近似为线性缓变结;(b)如果空间电荷区宽度远大于扩散长度 $2\sqrt{Dt}$,扩散结可近似为"单边"突变结

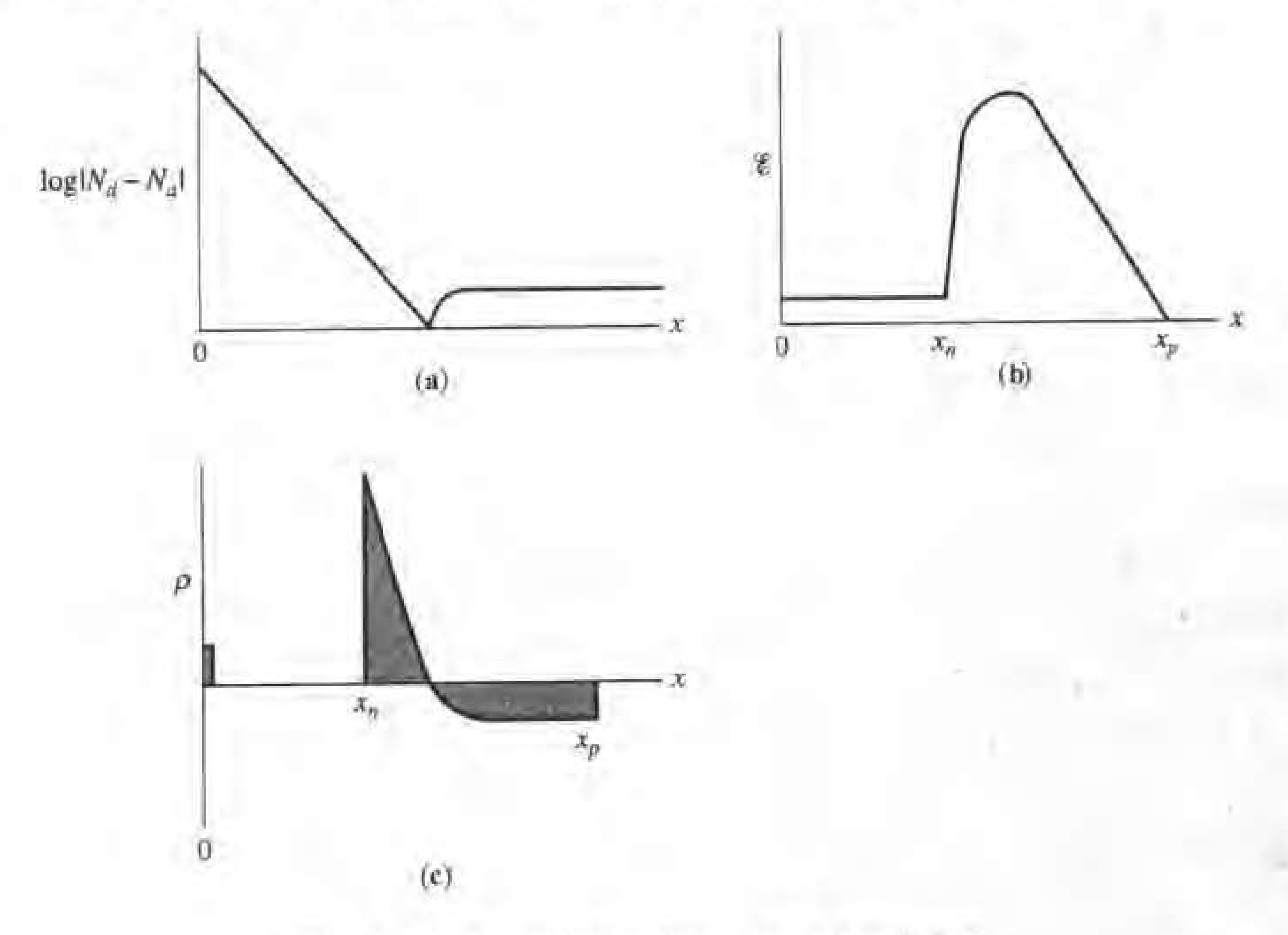


图 4.8　指数分布结的特性:(a)净掺杂浓度分布(半对数坐标);(b)电场分布;(c)空间电荷分布

4.2.3　异质结

现代半导体器件中,为了改善器件性能,可以用两种不同的半导体材料形成结。在继续讨论二极管和双极晶体管之前,我们先简要地讨论一下异质结的性质,更深入的分析已经超出了本书的范围[2]。在第 8 章讨论 Si-SiO_2 界面时,还会用到这些概念。

为分析异质结系统,考虑两种半导体材料 1 和 2,它们各自的组分和掺杂浓度均匀。与图 4.3 所示的同质结类似,画出两块半导体材料空间分离时的能带图,如图 4.9(a)所示。然而,

异质结中的 $E_{g1} \neq E_{g2}$，两种材料的电子亲和能 X 也可能不相同。假定材料 1 是 n 型，掺杂浓度为 N_{d1}，带隙为 E_{g1}，电子亲和能为 χ_1；材料 2 是 p 型，掺杂浓度为 N_{a2}，带隙为 E_{g2}，电子亲和能为 χ_2。

我们只考虑图 4.9(a)中的"跨立型"异质结，宽带隙材料的导带边比窄带隙的具有更高的能量，而价带边的能量则更低。尽管高低结(两种材料具有相同的导电类型)具有一些有趣的性质，但是这里我们只考虑 pn 异质结。特别是假定异质结左侧是宽带隙 n 型材料，右侧是窄带隙 p 型材料。这种结合方式对重要的异质结双极晶体管尤其有用，异质结双极晶体管将在第 6 章和第 7 章中讨论。

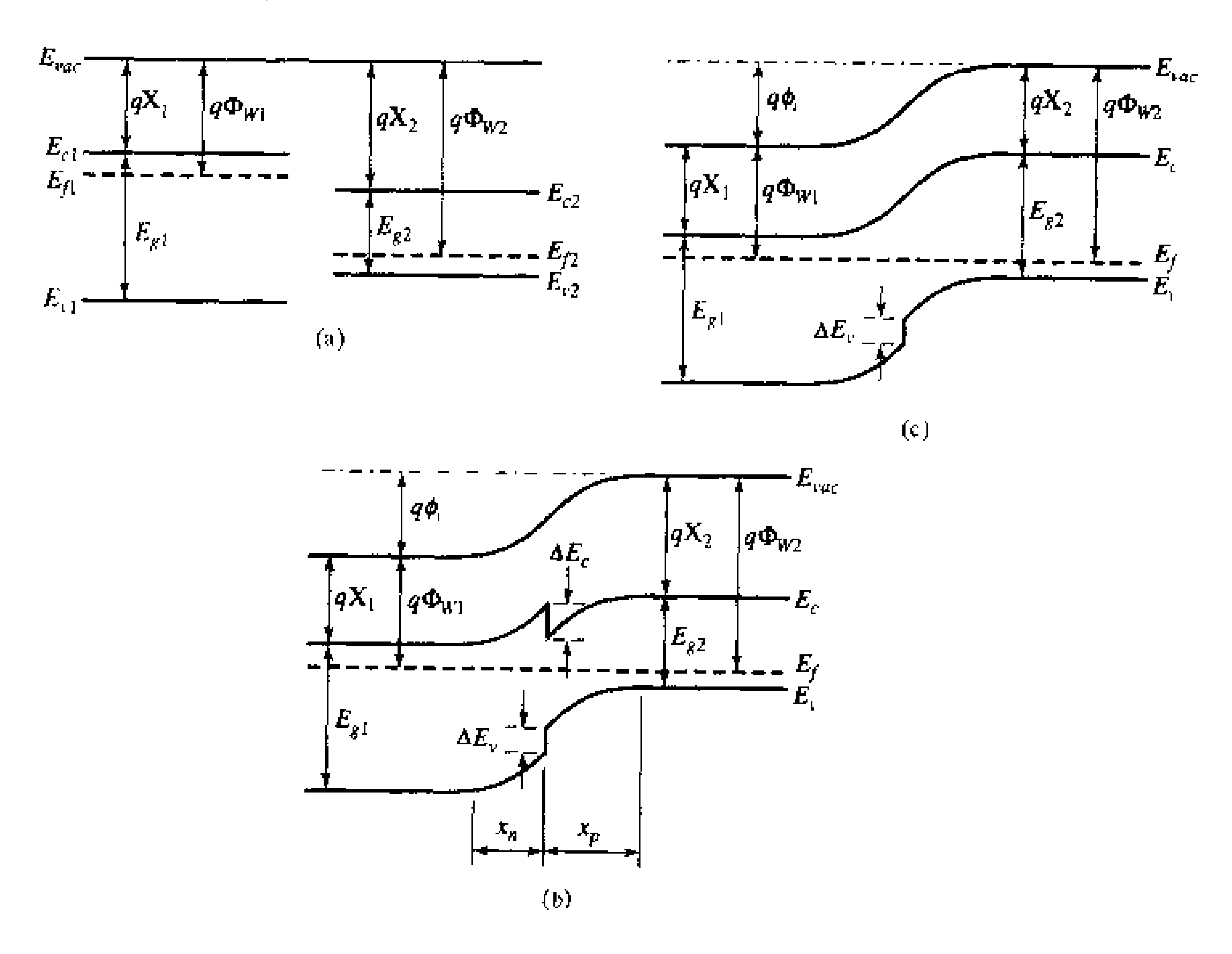


图 4.9 异质结的能带图：(a)空间分离的宽带隙 n 型材料和窄带隙 p 型材料；(b)当两种材料密切接触后，电荷移动达到系统平衡。真空能级保持连续，但是导带边和价带边可能不连续；(c) $\chi_1 = \chi_2$时的能带图

在以下分析中，将利用同质 pn 结的结果。假定两种半导体材料密切接触，并且两种材料的结上没有表面态存在。类似于同质结，当两种材料密切接触时，如图 4.9(b)所示，将会发生电荷流动以建立平衡，形成的电场将阻止电荷向低浓度区扩散。达到平衡时，整个系统的 Fermi能级一致。然而，两种材料的导带边 E_c和价带边 E_v，在分界处不一定连续。因为 $E_{g1} \neq E_{g2}$，至少有一个带边是不连续的。在界面处两种材料带边的相对位置取决于电子亲和能和禁带宽度的大小。

这时，导带上将出现"尖峰"和"凹槽"。这些特性取决于 ΔE_c 的大小以及两种半导体的掺杂。例如，自由电荷会在凹槽中积累，使器件的直流和交流性能变差。图 4.9(b)中，凹槽底部高于 Fermi 能级，这样只有少量的电荷被存储在那里，下面的讨论我们将其忽略。

假设真空能级在整个结中是连续的,那么真空能级的变化,即我们定义的内建电势 ϕ_i,能够从结两侧中性区的材料参数得到,其值等于两种半导体的功函数 Φ_W之差

$$\Phi_i = \Phi_{W2} - \Phi_{W1} = \left(\frac{X_2 + E_{c2} - E_f}{q}\right) - \left(\frac{X_1 + E_{c1} - E_f}{q}\right) \tag{4.2.17}$$

其中

$$E_{c1} - E_f = kT\ln\left(\frac{N_{c1}}{N_{d1}}\right) \tag{4.2.18}$$

和

$$E_{c2} - E_f = E_{g2} - (E_f - E_{v2}) = E_{g2} - kT\ln\left(\frac{N_{v2}}{N_{a2}}\right) \tag{4.2.19}$$

那么,

$$\begin{aligned}\phi_i &= X_2 - X_1 + \frac{E_{g2}}{q} - \frac{kT}{q}\ln\left(\frac{N_{v2}}{N_{a2}}\right) - \frac{kT}{q}\ln\left(\frac{N_{c1}}{N_{d1}}\right) \\ &= X_2 - X_1 + \frac{E_{g2}}{q} - \frac{kT}{q}\ln\left(\frac{N_{c1}N_{v2}}{N_{d1}N_{a2}}\right)\end{aligned} \tag{4.2.20}$$

分析中要注意结两侧材料的两种性质发生了变化:带隙(与每种半导体中少数载流子和多数载流子性质有关)和电子亲和能(与整个结的性质有关)。为了理解这两个因素的影响,下面将逐一讨论。先假定两种材料的电子亲和能相等($X_1 = X_2$),这样导带是连续的,如图 4.9(c)所示。因为 $E_{g1} \neq E_{g2}$,则价带必然不连续。

在 $X_1 = X_2$的条件下,等式(4.2.20)变为

$$\begin{aligned}\phi_i &= \frac{E_{g2}}{q} - \frac{kT}{q}\ln\frac{N_{v2}}{N_{a2}} - \frac{kT}{q}\ln\frac{N_{c1}}{N_{d1}} \\ &= \frac{E_{g2}}{q} - \frac{kT}{q}\ln\left(\frac{N_{c1}N_{v2}}{N_{d1}N_{a2}}\right)\end{aligned} \tag{4.2.21}$$

这样,异质结的内建电势与同质结具有相同的表达式,只是表达式中的带隙为窄带隙材料 2 的带隙(考虑一种特殊情况,假定 n 型材料的带隙大于 p 型材料)。

因为 $X_1 = X_2$,所以 $\phi_i = \phi_{Bn}$,这时形成了阻止电子从宽带隙材料 1 向窄带隙材料 2 扩散的势垒。无论是直接观察图 4.9(c),还是采用类似的推导,我们都能得出阻止空穴从材料 2 向材料 1 扩散的势垒

$$\phi_{Bp} = \frac{E_{g1}}{q} - \frac{kT}{q}\ln\left(\frac{N_{c1}N_{v2}}{N_{d1}N_{a2}}\right) \tag{4.2.22}$$

这个表达式与由材料 1 组成的同质结的表达式类似。根据等式(4.2.22),$\phi_{Bp} = \phi_{Bn} + E_{g1} - E_{g2} = \phi_{Bn} + \Delta E_g$。在第 5 章讨论异质结电流,以及第 6 章讨论异质结双极晶体管时会发现,多数载流子和少数载流子运动过程中遇到的势垒高度不同。

现在考虑两种材料的电子亲和能不同带来的变化。假定其他的材料性质不变,这样我们就可以只关注电子亲和能的改变所带来的变化。尽管两种材料的电子亲和能不同,仍可假定界面处的真空能级连续,这时导带边必定不连续,导带边能量差 $\Delta E_c = X_2 - X_1$。一般价带边能

量差也随之改变。(注意符号 ΔE_c 代表电子亲和能之差,所以它就是界面处导带边的不连续值。$E_{c2}-E_{c1}$ 是远离界面的中性区的导带边能量差。)

界面处能带不连续将影响到附近的空间电荷区。在材料的中性区,导带边和 Fermi 能级的距离取决于材料的掺杂浓度(等式(4.2.18)和(4.2.19)),所以,两种材料中性区的导带边之差并不取决于电子亲和势之差,而是取决于等式(4.2.18)和(4.2.19),即使 $X_1 \neq X_2$。

$$E_{c2} - E_{c1} = E_{g2} - kT\ln\left(\frac{N_{c1}N_{v2}}{N_{d1}N_{a2}}\right) \tag{4.2.23}$$

由于两边中性区的 $E_{c2}-E_{c1}$ 不变,并且界面处 ΔE_c 非零,则总的能带弯曲必定随 ΔE_c 而增加(图 4.9(b)是 $\Delta E_c>0$ 的情况)。因为 Φ_{w1} 或 Φ_{w2} 发生变化,当 $\Delta E_c>0$ 时,真空能级的改变(定义为内建电势 ϕ_i)也会随之增加。由于异质结带边不连续,内建电势不再等于两种材料远离结处的导带边能量之差,因此,与同质结相比,内建电势的概念并不常用。

当 $X_1 \neq X_2$ 时,能带弯曲增大,表明与 $X_1=X_2$ 的异质结相比,这时的耗尽区更宽,如图 4.9(b)所示。耗尽区宽度可从 Poisson 方程求出,在各种半导体材料中的分析与同质结一样,区别只是界面处的边界条件不同。由 Gauss 定理可求出每种材料耗尽区的电场分布,得到的解与等式(4.2.4)和(4.2.5)相同,只是两种材料的介电常数不同,分别表示为 ϵ_1 和 ϵ_2。界面处 $\mathscr{D}$ 连续(假定界面处不存在面电荷)

$$\epsilon_1\mathscr{E}_1(0) = \epsilon_2\mathscr{E}_2(0) \tag{4.2.24}$$

因此,当 $\epsilon_1 \neq \epsilon_2$ 时,界面处电场不再连续。令 x_n 和 x_p 为结两侧耗尽区宽度,然后,将表达式 $\mathscr{E}_1(0)=(q/\epsilon_1)N_dx_n$ 和 $\mathscr{E}_2(0)=(q/\epsilon_2)N_ax_p$ 代入等式(4.2.24),得到与同质结相同的等式 $N_dx_n=N_ax_p$(电中性条件)。

与同质结一样,通过对异质结两侧的电场积分,再加上结两侧电势的变化,就可以得到电势分布

$$\phi_1(-\infty) - \phi_1(x) = \frac{q}{2\epsilon_1}N_d(|x| - x_n)^2 \tag{4.2.25}$$

和

$$\phi_2(x) - \phi_2(\infty) = \frac{q}{2\epsilon_2}N_a(x_p - x)^2 \tag{4.2.26}$$

利用 ϕ_i 的定义可知两种材料能带弯曲之和就是 ϕ_i。从

$$\phi_i = \phi_1(-\infty) - \phi_2(\infty) \tag{4.2.27}$$

和

$$\phi_1(0) = \phi_2(0) \tag{4.2.28}$$

得到

$$\phi_i = \frac{qN_dx_n^2}{2\epsilon_1} + \frac{qN_ax_p^2}{2\epsilon_2} \tag{4.2.29}$$

正如前面指出的,当 $\Delta E_c>0$ 时 ϕ_i 增加,表明耗尽区宽度也增加。

求解结两侧的电压降,得到

$$\phi_{i1} = \phi_i \frac{\epsilon_2 N_a}{\epsilon_1 N_d + \epsilon_2 N_a} \tag{4.2.30}$$

和

$$\phi_{i2} = \phi_i \frac{\epsilon_1 N_d}{\epsilon_1 N_d + \epsilon_2 N_a} \tag{4.2.31}$$

与同质结一样，内建电势主要降落在轻掺杂材料一侧，并受到不同的介电常数的影响。然而，对异质结来说，研究降落在各半导体上的内建电势之比比同质结具有更重要的意义。因为带边 E_c 和 E_v 不连续，出现如图 4.9(b) 所示的尖峰，阻止载流子流过结区。尖峰作为势垒阻止电子从材料 1 流向材料 2，其大小取决于降落在每一侧的电势之比。如果材料 2 中的受主浓度比材料 1 中的施主浓度低，能带弯曲主要发生在材料 2 中，尖峰的最高点比材料 2 的中性区导带边低，如图 4.10(a) 所示。阻止电子从材料 1 流向材料 2 的势垒 ϕ_{Bn} 恰好等于两种材料的中性区导带边之差。然而，如果材料 2 的掺杂浓度比材料 1 高，能带弯曲主要发生在材料 1 中，如图 4.10(b) 所示，此时尖峰的最高点将比材料 2 的中性区导带边高。在这种情况下，阻止电子从材料 1 流向材料 2 漂移的势垒 ϕ_{Bn} 比两种材料的中性区导带边之差大，任何包含这种结的器件的工作原理将会发生重大变化。在更细致的分析中，必须考虑邻近尖峰的凹槽的影响，因为自由载流子会在那里积累，影响器件的性能。

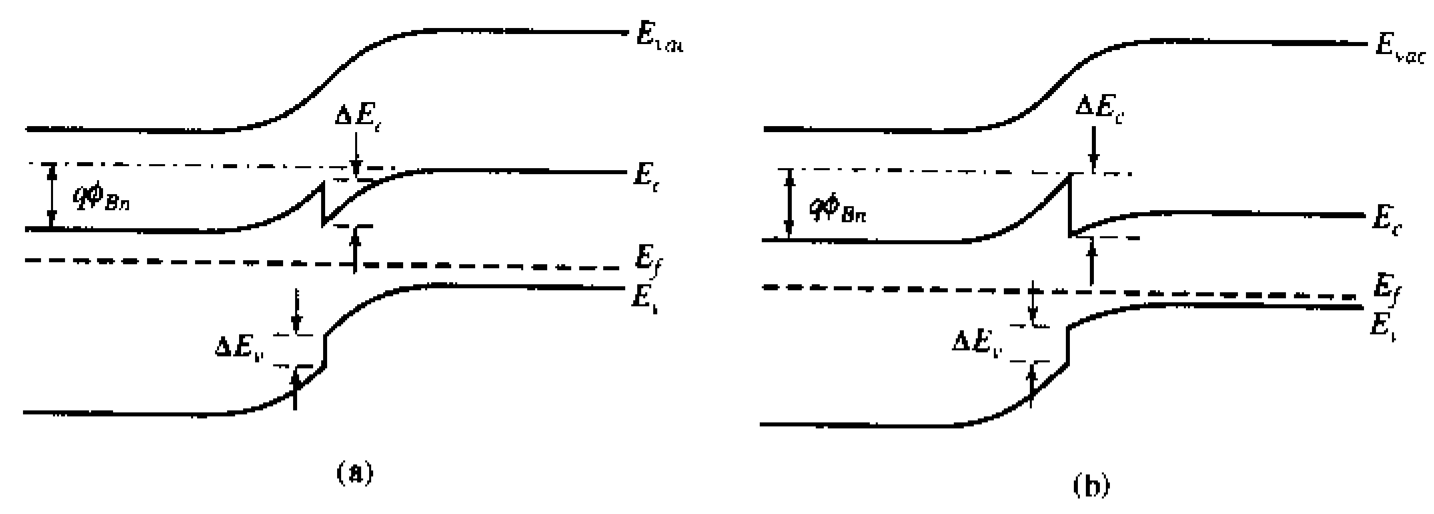


图 4.10　电子从 n 型半导体注入到窄带隙 p 型半导体遇到的势垒取决于导带边能量差以及 p 型区的掺杂。(a) p 型区轻掺杂；(b) p 型区重掺杂，导带尖峰阻止电子注入

4.3　反偏 pn 结

4.2 节采用耗尽近似简化了热平衡下 pn 结的电学特性的分析。为了讨论外加偏压下的结的特性，再次采用耗尽近似，并利用一些关于外加偏压的假设。假定 p 型和 n 型区的电极是欧姆接触，可忽略电极上的压降。同时假定降落在中性区上的电压很小，并且 n 型区接地，电压 V_a 加在 p 型区上。在以上条件下，便可求出加在整个结上的总的外加电压。此外，在这些假设下，上节得到的热平衡条件下 Poisson 方程的解同样适用于偏压条件，只需将结上的总电势差从内建电势 ϕ_i 变为 $\phi_i - V_a$。

如果 V_a 为正，结上多数载流子运动遇到的势垒降低，耗尽区变窄。关于耗尽区宽度减小的一种直观的解释是：外加电压迫使多数载流子流向耗尽区边界，中和了部分空间电荷，这样就减小了整个耗尽区宽度。$V_a > 0$ 时，结上总的势垒是 $\phi_i - V_a$，结正偏。在正向偏置下，即使

V_a很小,也会有很大的电流流过。关于正向偏置的进一步讨论见第 5 章。

如果 p 型区加反向偏压,多数载流子运动遇到的势垒增高。同样,总的结上势垒仍可以表示成 $\phi_i - V_a$,不过这时 V_a是负值,结处于反偏。反偏时,多数载流子被外加电压从耗尽区边缘拉出,所以耗尽区变宽,这时,只有非常小的电流流过,因为偏压导致电子从 p 型区流向 n 型区,空穴从 n 型区流向 p 型区,而它们在各自的区域内是少数载流子,浓度很低。我们将在第 5 章中讨论反偏 pn 结的电流。

耗尽区宽度,最大电场 突变 pn 结的耗尽区宽度是电压的函数,将等式(4.2.11)中的内建电势 ϕ_i替换成 $\phi_i - V_a$,这样,

$$x_d = x_n + x_p = \left[\frac{2\epsilon_s}{q}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)(\phi_i - V_a)\right]^{1/2} \tag{4.3.1}$$

当 V_a的值远大于 ϕ_i时,突变结的耗尽区宽度与反偏电压的平方根成正比。对其他的掺杂分布,x_d的表达式可以通过修正热平衡的结论得到。例如,掺杂梯度为 a(等式(4.2.15))的线性缓变结的耗尽区宽度可由下式给出

$$x_d = \left[\frac{12\epsilon_s(\phi_i - V_a)}{qa}\right]^{1/3} \tag{4.3.2}$$

通常结上最大场强 $\mathscr{E}_{max}$及其与外加电压的关系非常重要。突变结的情况很简单,因为电场随位置线性变化(等式(4.2.4)和(4.2.5)),所以,电场分布曲线下覆盖的面积代表了电势,可以写成最大场强的一半乘以耗尽层宽度

$$\frac{1}{2}\mathscr{E}_{max}x_d = (\phi_i - V_a)$$

于是,突变结中

$$\mathscr{E}_{max} = \frac{2(\phi_i - V_a)}{x_d} \tag{4.3.3}$$

对线性缓变结,可以用习题 4.5 的结论求出最大场强为

$$\mathscr{E}_{max} = \frac{3(\phi_i - V_a)}{2x_d} \tag{4.3.4}$$

电容 3.2 节中分析了金属-半导体结中的电容效应,它来源于外加电压变化导致存储在结耗尽区中的电荷的变化。pn 结有类似的现象,只是耗尽区沿两个方向随偏压变化。通常,对掺杂浓度分别为 N_a和 N_d的突变结来说,根据结耗尽区任意一侧电荷 Q_s(单位面积)的表达式,就可以求出单位面积的小信号电容

$$Q_s = qN_dx_n = qN_ax_p \tag{4.3.5}$$

根据小信号电容 C(单位面积)的定义

$$C = \frac{dQ}{dV_a} = qN_d\frac{dx_n}{dV_a} = qN_a\frac{dx_p}{dV_a} \tag{4.3.6}$$

考虑 $x_p = (N_d/N_a)x_n$和 $x_d = x_n + x_p$,从等式(4.3.1)得到

$$\frac{dx_n}{dV_a}=\frac{1}{N_d}\left[\frac{\epsilon_s}{2q\left(\frac{1}{N_a}+\frac{1}{N_d}\right)(\phi_i-V_a)}\right]^{1/2} \tag{4.3.7}$$

和

$$C=\left[\frac{q\epsilon_s}{2\left(\frac{1}{N_a}+\frac{1}{N_d}\right)(\phi_i-V_a)}\right]^{1/2} \tag{4.3.8}$$

这样,当$|V_a|$远大于ϕ_i时,突变pn结的电容近似与反偏电压的平方根成反比。将等式(4.3.1)代入(4.3.8)中,得到$C=\epsilon_s/x_d$,这是小信号电容的一般表达式。

当pn结一侧的掺杂浓度增高时,该侧的耗尽层宽度变窄。高掺杂一侧的耗尽区宽度只占整个空间电荷区宽度的一小部分,一般可忽略。此时,pn结的电场分布和空间电荷分布与具有整流作用的金属-半导体接触十分相似。[4]

尽管第3章已给出了电容的一般表达式$C=\epsilon_s/x_d$,由于其重要性,这里再仔细推导一遍以示强调。考虑任意掺杂的pn结,耗尽区从$-x_p$延伸到x_n,如图4.11(a)所示。储存在单位面积下x与耗尽区边界x_n之间的电荷Q为

$$Q=q\int_x^{x_n}N\,dx \tag{4.3.9}$$

其中净掺杂浓度$N=N_d-N_a$。考虑到$\mathscr{E}_x(x_n)=0$,从Gauss定理可求出x点的电场,如图4.11(b)所示

$$-\mathscr{E}_x(x)=\frac{1}{\epsilon_s}\int_x^{x_n}qN\,dx=\frac{Q}{\epsilon_s} \tag{4.3.10}$$

当外加电压V_a改变一小量dV_a时,n型区一侧的耗尽区宽度将随之改变dx_n,储存在x与x_n之间的电荷改变量是

$$dQ=qN(x_n)dx_n \tag{4.3.11}$$

因此,当电压变化时x点的电场改变量为

$$-d\mathscr{E}_x=\frac{dQ}{\epsilon_s}=\frac{q}{\epsilon_s}N(x_n)\,dx_n \tag{4.3.12}$$

$\mathscr{E}_x\sim x$曲线下的面积对应于总的电势ϕ_i-V_a,电压的微分就对应于曲线下面积的改变,如图4.11(b)的阴影部分所示

$$dV_a\approx -x_d\,d\mathscr{E}_x=\frac{x_d}{\epsilon_s}dQ \tag{4.3.13}$$

将等式(4.3.13)代入小信号电容的定义(等式(4.3.6)),得到

$$C=\frac{dQ}{dV_a}=\frac{\epsilon_s}{x_d} \tag{4.3.14}$$

---

[4] 换个角度考虑,当电荷分布越来越集中,并最终接近于δ函数时,电场分布非常陡峭。这时的重掺杂半导体可看做是理想金属,所有电荷都集中在表面。

这就证明了对任意掺杂的结,这一简单的关系式成立。x_d与电压的关系以及电容与电压的关系都和掺杂分布有关。

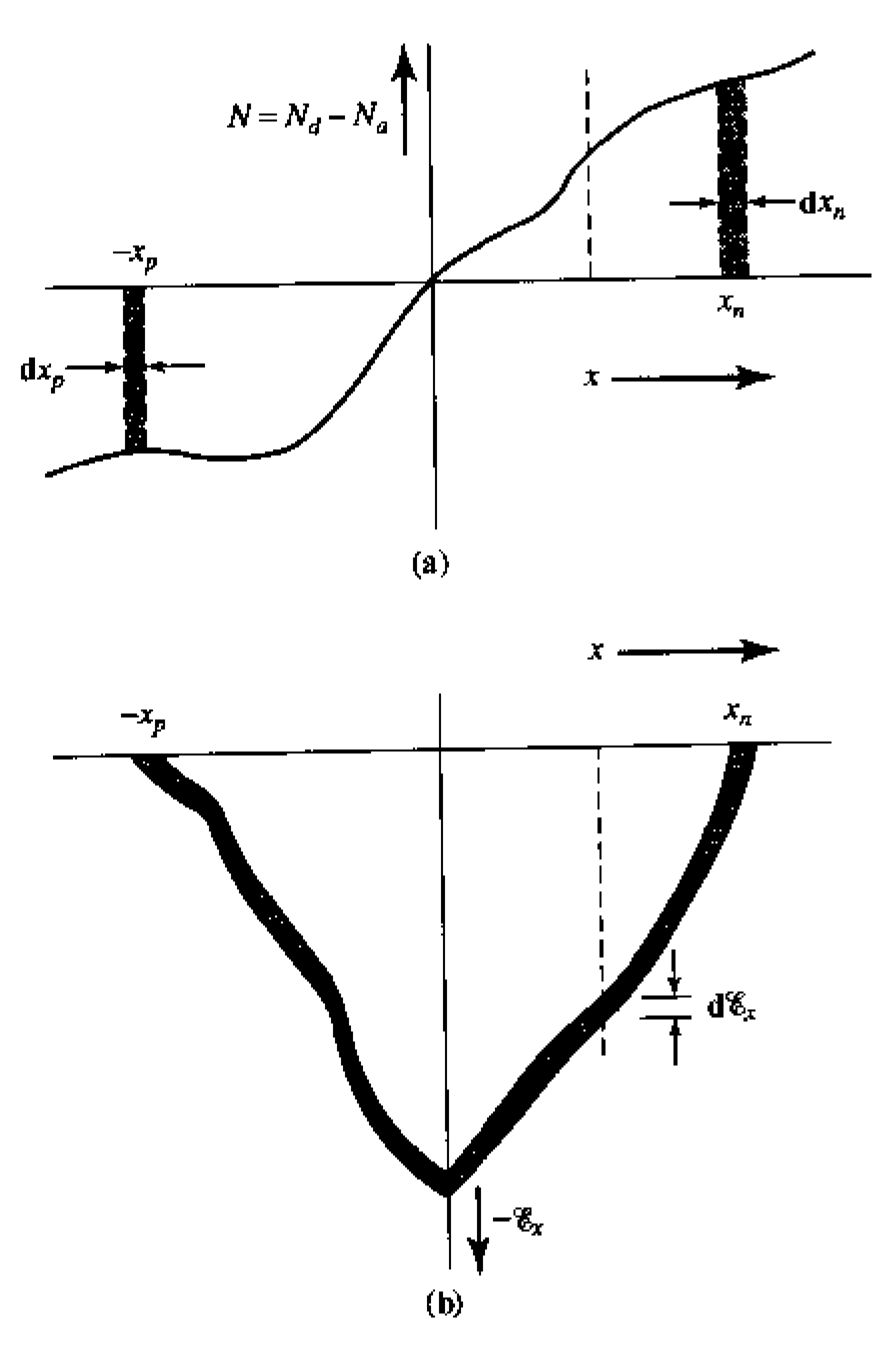


图4.11 任意掺杂pn结。(a)外加电压对空间电荷区边界处载流子浓度的调制;(b)外加电压的微小变化导致的电场分布的改变

与金属-半导体结类似,通过测量pn结小信号电容随外加电压的变化,可以确定掺杂浓度的分布。但是pn结的情况要复杂一些,因为在$N=N_d-N_a=0$的结平面处,耗尽区向两个方向延伸。为了更深入地研究这个问题,对等式(4.3.14)求微分,同时考虑$x_d=x_n+x_p$,得到

$$\frac{dC}{dx_n}=-\frac{\epsilon_s}{(x_n+x_p)^2}\left(1+\frac{dx_p}{dx_n}\right) \tag{4.3.15}$$

因为结两侧的电荷变化量相等,所以

$$|dQ|=|qN(-x_p)dx_p|=qN(x_n)dx_n \tag{4.3.16}$$

并且

$$\frac{dC}{dx_n}=-\frac{C^2}{\epsilon_s}\left(1+\frac{N(x_n)}{N(-x_p)}\right) \tag{4.3.17}$$

将这个表达式与等式(4.3.11)及电容的定义结合起来,得到

$$N(x_n) = -\frac{C^3}{\epsilon_s q(\mathrm{d}C/\mathrm{d}V_a)}\left(1 + \frac{N(x_n)}{N(-x_p)}\right) \tag{4.3.18}$$

如果结的 p 型区掺杂远高于 n 型区,等式(4.3.18)右边的因子变为 1,等式也将简化成金属-半导体结的等式(3.2.10)。这在物理上是合理的,因为耗尽区很难在高掺杂半导体中延伸,正如在金属-半导体结中金属一侧的耗尽区宽度可以忽略。然而,更一般的情况是 $N(x_n)$ 的表达式中包含因子$[1+N(x_n)/N(-x_p)]$,使得通过测量电容来确定掺杂分布变得复杂,必须采用迭代技术或者测量其他参量。事实上,等式(4.3.18)中的微分 $\mathrm{d}C/\mathrm{d}V_a$项增加了确定掺杂分布的复杂性。对有噪声的实验信号微分,需要实验设备具有更高的屏蔽性,以及合适的数值平均技术。

对于异质结电容,考虑到两侧的介电常数和内建电势的不同,修正同质结的电容表达式,得到(类似于等式(4.3.18))

$$\frac{1}{C} = \frac{x_n}{\epsilon_1} + \frac{x_p}{\epsilon_2} = \sqrt{\frac{2\phi_i}{q}\frac{\epsilon_1 N_d + \epsilon_2 N_a}{\epsilon_1 N_d \epsilon_2 N_a}} \tag{4.3.19}$$

发光二极管 1.1 节讨论了光电导,即光照射在半导体上激发出的光生载流子导致半导体的电导率增加。当存在 pn 结时,光与半导体之间的相互作用以其他的方式表现出来。pn 结耗尽区电场在电子-空穴对复合前将它们分离,并加速光生载流子向中性区的运动,这样就对电流有贡献。其中空穴向 p 型区运动,电子向 n 型区运动。

如果没有外加偏压,中性区的过剩载流子将使 Fermi 能级向带隙中部移动,使得 pn 结上的电势低于热平衡内建电势,产生光电流和光生电压。这一光生伏特工作模式能够提供能源,而且是直接将光能转化为电能的太阳能电池的基础。

当反偏 pn 结被光照射时,光致载流子导致反偏电流激增。没有光照时,耗尽区中载流子很少,因而反偏电流很小。当入射光激发出电子-空穴对时,耗尽区载流子数量激增,反向电流也相应地增加。反偏结中光电流和暗电流的比值可以相差好几个数量级。如果 pn 结二极管的厚度小于光的透入深度或者二者可以相比拟时,尽管整个二极管中都可产生光生载流子,但是当载流子产生发生在耗尽区时,收集效率最高。耗尽区的高电场使光生电子-空穴对在复合之前被分离,并且加速了载流子向中性区运动,在那里它们是多数载流子,对电流有贡献。外加电压同时增加了耗尽区宽度,使得入射光产生更多的光致电流。

为了增加效率,耗尽区应当越宽越好。特制的光电二极管中,在重掺杂的 p 型区和 n 型区之间增加一层近乎本征(i)的轻掺杂层,构成 p-i-n 二极管结构。只要在二极管两端加适当的反偏电压,这一 i 层很容易被耗尽,这样载流子的有效收集区变大,相应光致电流也增大。因为载流子在耗尽区被电场加速,p-i-n 二极管对光强的变化响应速度很快。相反,中性区产生的少数载流子向收集区扩散的过程非常缓慢。尽管 p-i-n 二极管效率高、速度快,但增加一层轻掺杂层并不容易,因为与传统的集成电路工艺不易兼容。因此,p-i-n 二极管仅用于特殊场合,而常规的光探测则采用能够与其他功能电路集成的 pn 结二极管。

4.4 结的击穿

上节中看到,pn 结的耗尽区宽度和最大场强随反偏电压的增加而增加。凭直觉我们知道

这种增加必定有物理上的极限。高电压下,构成器件的某些材料,比如二氧化硅绝缘层,或者封装材料可能被损坏,或者通过 pn 结的电流会出现快速增大。第一种情况通常不可恢复,器件被彻底损坏。第二种情况有重要的实际意义,即结中的电流导致势垒击穿,一般不是破坏性的(除非大电流使部分结熔化)。结的击穿电压与 pn 结结构以及掺杂分布有明确的关系,因此,结的击穿特性可以预测。

高电场下,半导体可能发生两种类型的击穿。一种是 Zener 击穿,指的是电场强到足以将被共价束缚的电子释放出来,产生对电流有贡献的两个载流子——电子和空穴。用能带图描述,就是电子直接从价带跃迁到导带,没有与其他粒子相互作用。这类击穿涉及了 3.4 节讨论欧姆接触时介绍的电子隧穿通过能量势垒的现象。另一种是雪崩击穿。自由载流子在两次碰撞之间从电场获得足够的能量,与晶格发生碰撞时能打破共价键的束缚,产生电子-空穴对。如果电场足够高,所有的载流子以及新产生的电子空穴对都能从电场中获得能量,参与随后的雪崩碰撞,导致空间电荷区载流子的急剧增加。

在给定的 pn 结中,两种击穿方式中的某一种会起主导作用。在详细讨论了两种击穿过程之后,我们就会搞清楚决定击穿机构的因素。下面讨论最普遍的雪崩击穿,然后考虑 Zener 击穿,最后比较这两种过程。

4.4.1 雪崩击穿*

考虑反偏 pn 结空间电荷区中电子的运动。在与晶格原子碰撞失去能量前,电子移动的平均距离称平均自由程 λ。两次碰撞之间从电场 $\mathscr{E}$ 获得的能量 ΔE 为

$$\Delta E = q\int_0^\lambda \boldsymbol{\mathscr{E}} \cdot \mathrm{d}\boldsymbol{x} \tag{4.4.1}$$

等式(4.4.1)中的黑体字表示向量积。如果电子在与原子碰撞前从电场获得足够的能量,在碰撞时就能打破原子核对一个电子的束缚,碰撞后该区域出现三个自由载流子——初始电子和碰撞产生的电子-空穴对,如图 4.12 所示。假定这三个载流子的质量相同,能量和动量守恒要求初始电子的动能至少是 $3E_g/2$,才能打破原子键(习题 4.11)。

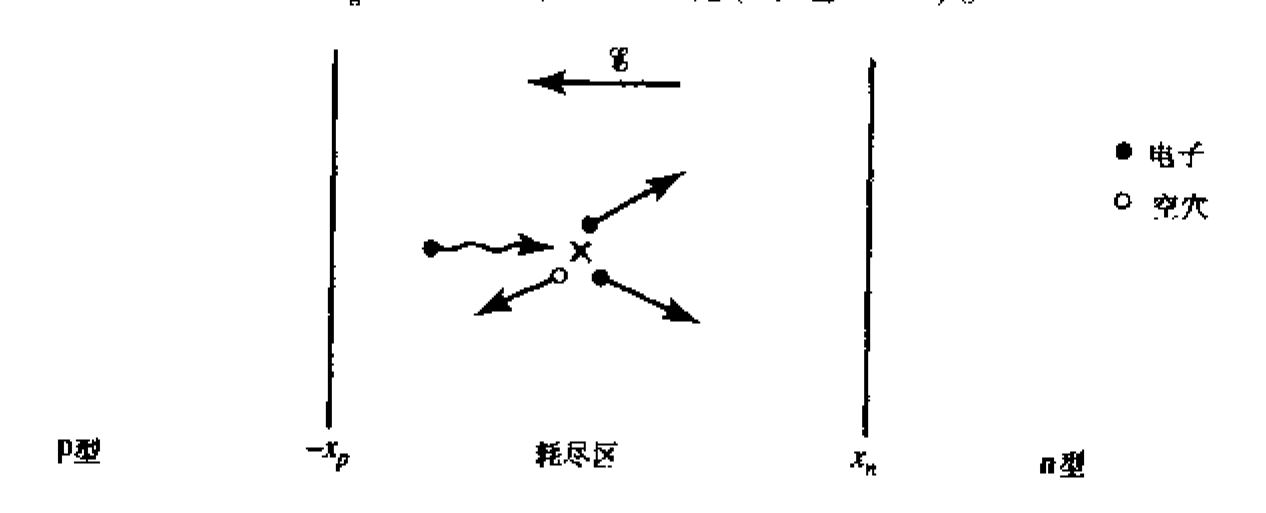


图 4.12 雪崩击穿过程示意图。一个入射电子(波浪箭头)从电场中获得足够的能量,在与晶格碰撞时,从硅-硅键中激发出一个电子,产生电子-空穴对

为简单起见,假设突变结 n 型区掺杂远高于 p 型区。在这种情况下,正如第 5 章将要讨论的,在中等的反向电压下,进入耗尽区的大部分载流子是 p 型区的电子。分析中可忽略从 n 型区进入耗尽区的空穴。空间电荷区边界处的电场较低,几乎没有载流子在与晶格中原子碰撞

而失去动能前，从电场获得足够的能量来产生电子-空穴对。因此，雪崩击穿一般都集中在电场较高的空间电荷区中心部分。这一中心区域在图4.13(a)所示的电场分布图上被标注为x_1。

现在考虑x_1区域位于x处的dx范围内由于雪崩所产生的载流子数，如图4.13(b)所示。假设从左方x_a处进入x_1区域的电子浓度为n_0。x_a到x之间的雪崩过程使这一浓度增加，这样从左侧进入x处体积为Adx区域的浓度就是n_0+n_1。在通过dx区域时，电子产生电子-空穴对的几率由电离率α_n与长度dx的乘积决定。当电场很强时，电子获得能量的速度非常快，可以认为电离率是电场强度的函数，因而也是位置的函数。在dx处从左侧进入的电子造成的电子浓度增量(和空穴密度增量)为

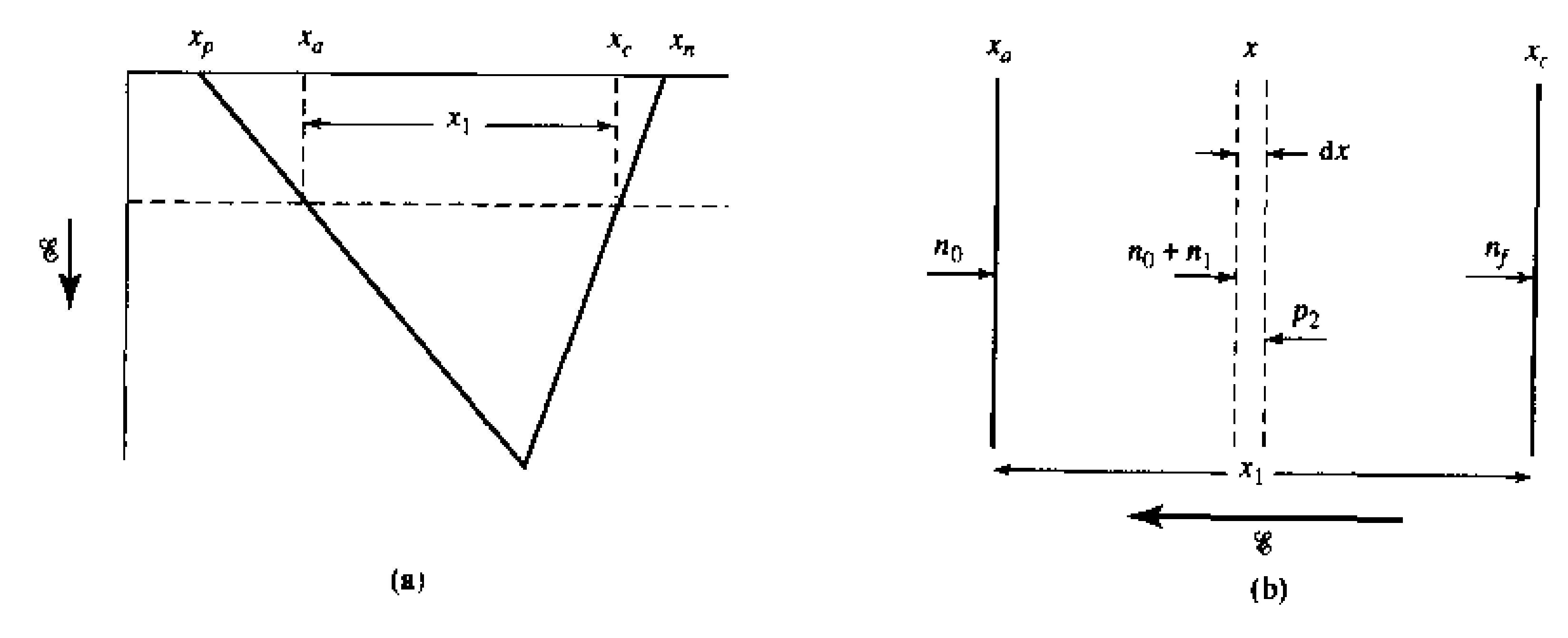


图4.13 (a)碰撞电离发生在空间电荷区高场区x_1;(b)从左侧流入的电子和从右侧流入的空穴在x点处dx范围内产生载流子对

$$\mathrm{d}n' = \mathrm{d}p = \alpha_n n\,\mathrm{d}x = \alpha_n(n_0 + n_1)\mathrm{d}x \tag{4.4.2}$$

因为假设没有空穴从x_c进入电离区，在x处从右侧进入无限小空间的空穴都是在x到x_c之间产生的，假定这些空穴浓度为p_2。在dx中，空穴也是雪崩倍增的，得到的空穴和电子浓度增量为

$$\mathrm{d}n'' = \mathrm{d}p = \alpha_p p\,\mathrm{d}x = \alpha_p p_2\,\mathrm{d}x \tag{4.4.3}$$

式中α_p是空穴的电离率。dx内总的电子浓度的增加等于(dn'+dn'')(等式(4.4.2)和(4.4.3))

$$\mathrm{d}n = \alpha_n(n_0 + n_1)\mathrm{d}x + \alpha_p p_2\,\mathrm{d}x \tag{4.4.4}$$

如果令n_f为到达x_c的电子浓度

$$n_f = n_0 + n_1 + n_2 \tag{4.4.5}$$

式中n_2是在x和x_c之间产生的电子浓度，因为电子和空穴成对产生，所以$n_2 = p_2$，于是可以写出

$$\frac{\mathrm{d}n}{\mathrm{d}x} = (\alpha_n - \alpha_p)(n_0 + n_1) + \alpha_p n_f \tag{4.4.6}$$

为进一步分析，必须得到关于电离率α_n和α_p的其他信息。如果令$\alpha_n = \alpha_p$，可以得到关于雪崩击穿一些比较有用的结论，虽然只是近似。假定$\alpha \equiv \alpha_n = \alpha_p$，对等式(4.4.6)积分，取边界条件

$n(x_a)=n_0$和$n(x_c)=n_f$,积分结果是

$$n_f - n_0 = n_f \int_{x_a}^{x_c} \alpha \, dx \tag{4.4.7}$$

定义n_f(离开空间电荷区的电子浓度)和n_0(进入空间电荷区的电子浓度)的比值为倍增因子M

$$M = \frac{n_f}{n_0} = \frac{1}{1 - \int_{x_a}^{x_c} \alpha \, dx} \tag{4.4.8}$$

当等式(4.4.8)中分母的积分接近1时,倍增因子将趋于无穷大。所以,雪崩击穿的条件也可以定义为

$$\int_{x_a}^{x_c} \alpha \, dx = 1 \tag{4.4.9}$$

如果没有$\alpha_n=\alpha_p$的条件,那么得到的积分式比等式(4.4.9)复杂得多,但是形式与等式(4.4.8)类似,并且令分母为零也可得到雪崩击穿的条件。

前面已经提到,电离率强烈依赖于电子能量,因为碰撞电离所需的能量是由电场提供给载流子的,因而电子能量又取决于电场。对缓变电场由下述讨论可得到关于电离率α的一个近似合理的表达式。x处的碰撞电离与n^*成正比,n^*是x处具有足够能量产生电子-空穴对的电子浓度。所以,浓度n^*等于总的电子浓度n乘以电子移动距离d并获得足够能量且没有发生碰撞的概率,即

$$n^* = n \exp\left(-\frac{d}{\lambda}\right) \tag{4.4.10}$$

式中λ是平均自由程。令E_1为产生碰撞电离所需的最小能量,$\mathscr{E}$为加速电子的平均电场,那么距离d可以从等式(4.4.1)得到

$$d = \frac{E_1}{q\mathscr{E}} \tag{4.4.11}$$

假设电子在获得足够的碰撞电离能量后马上发生碰撞,那么在距离dx中发生碰撞电离的数量和dx/d成正比,于是

$$dn = A'n^* \frac{dx}{d} = \frac{A'q\mathscr{E}}{E_1}\left[\exp\left(-\frac{E_1}{\lambda q\mathscr{E}}\right)\right] n \, dx \tag{4.4.12}$$

式中A'是比例系数。比较等式(4.4.2)和(4.4.12),可以得到电离率较为合理的表达式

$$\alpha = K\mathscr{E} \exp\left(-\frac{B}{\mathscr{E}}\right) \tag{4.4.13}$$

因为电离率强烈地依赖于电场,所以倍增因子也随电场快速增加。当所加电压接近击穿电压时,电场的微小增加将会导致电流的剧增。这一现象已在实际二极管中被证实。

不仅电离率随电场变化,是位置的函数,并且空间电荷区宽度也随电压变化。所以,估算等式(4.4.8)中的M很困难,经常采用的一个经验公式为

$$M = \frac{1}{1-(|V_R|/BV)^n} \quad (2 < n < 6) \tag{4.4.14}$$

式中 $V_R<0$ 是外加(反偏)电压,BV 是击穿电压,即电流迅速增长所对应的电压。

为了将击穿电压和材料参数相联系,考虑 $N_a \ll N_d$ 的单边突变结,假设最大场强达到临界值$\mathscr{E}_1$(使等式(4.4.8)中的积分接近于1的值)时结中发生击穿。最大电场由下式近似给出

$$\mathscr{E}_{max} = \left(\frac{2qN_a|V_R|}{\epsilon_s}\right)^{1/2} \tag{4.4.15}$$

击穿电压有如下的近似形式

$$BV = \frac{\epsilon_s \mathscr{E}_1^2}{2qN_a} \tag{4.4.16}$$

等式(4.4.16)表明掺杂浓度增加将使击穿电压下降,但是实际下降速度没有这么快。实际的二极管的击穿电压通常正比于 $N^{-2/3}$(N 为掺杂浓度)。击穿电压与掺杂浓度的依赖关系变缓的原因是,掺杂浓度降低导致耗尽区加宽,使能发生有效雪崩倍增的区域加宽。而掺杂浓度增加,所需的临界电场 $\mathscr{E}_1$ 略微高一些,如图4.14所示。

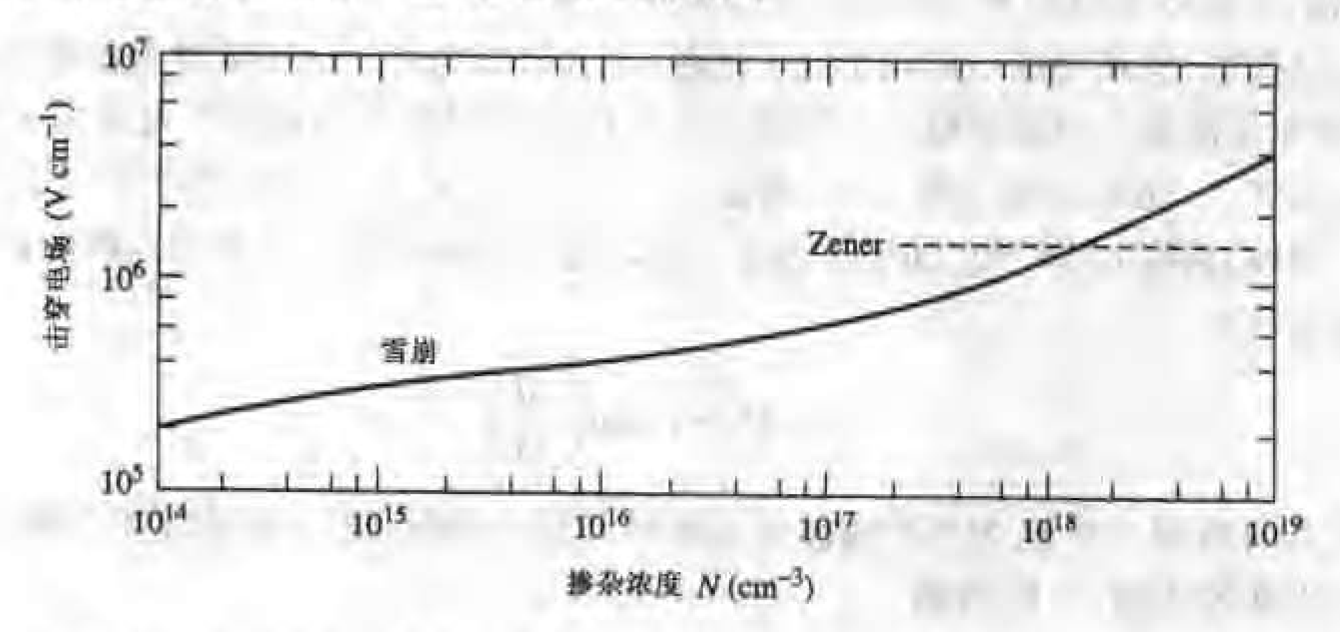


图4.14 硅器件中雪崩击穿和Zener击穿的临界电场与掺杂浓度的关系[3,4,5]

迄今为止的讨论只局限于一维情况,实际的平面结的有限的横向尺度将显著减小击穿电压。2.1节介绍了在二氧化硅的窗口上扩散形成的平面结。杂质垂直扩散进入硅,同时也在Si-SiO_2界面下横向扩散,形成了SiO_2边界下圆形的结,如图2.1(d)所示。这个拐角区的电场明显比结中其他区域高,导致低电压下就会发生击穿。曲率半径小的浅结,这种现象尤其严重。当结深 x_j 增加时会好一些。图4.15给出了硅单边突变平面结的结深对击穿电压的影响。

例题　pn结二极管的击穿

在电阻率 $\rho=1.5\Omega\cdot cm$ 硅衬底上掺杂硼制作二极管,加工工艺如下:

1. 表面磷原子浓度 $N'=5\times10^{15}cm^{-2}$

2. 1000℃推进扩散27分钟。

(a)如果在整个硅表面进行磷掺杂,击穿电压 BV 是多少?

(b)如果在衬底氧化层上的窗口区选择磷注入,形成隔离的二极管区域,那么击穿电压 BV 又是多少?

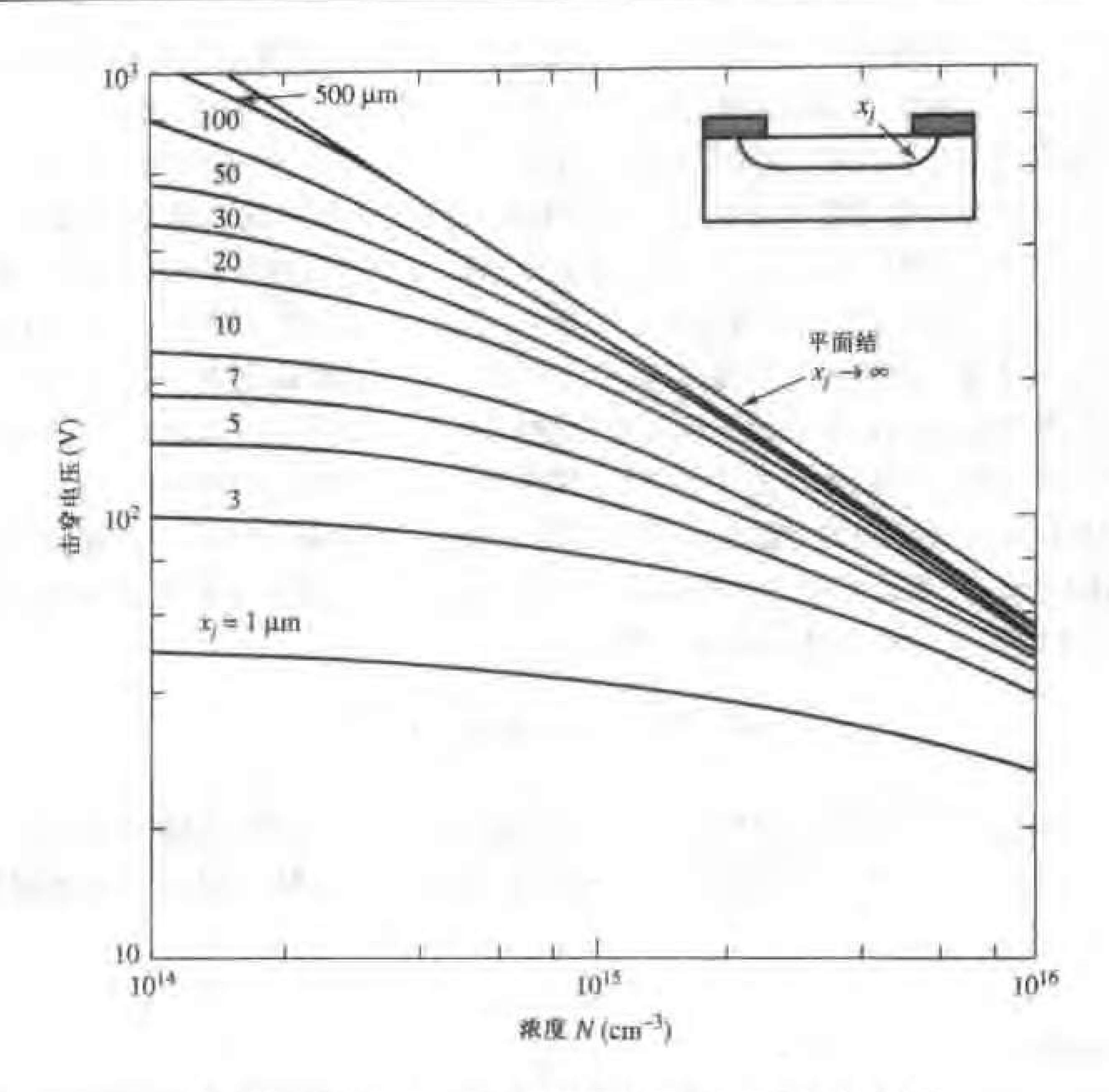


图4.15 硅单边突变平面结的击穿电压与结深的关系[6,7]

解:上述工艺可制备出 pn 结二极管。对于(a),形成一个平面结的单个平面二极管。对于(b),形成多个扩散区,每一个扩散区的横截面与图2.1(d)类似,所以(b)的二极管在拐角处的电场加强,因此它的击穿电压比(a)低。

恒定表面源的磷在1000℃下再分布27分钟,将形成 n 型杂质的 Gauss 分布。从图2.20知,磷在1000℃下的扩散系数是 $1.6\times10^{-13}\text{cm}^2\text{s}^{-1}$,所以 $\sqrt{Dt}=0.16\mu\text{m}$。

为了确定该二极管是否能被近似为单边结,必须确定磷再分布后的浓度分布。这种 Gauss 分布的表面浓度可以从等式2.5.13得到,取 $x=0$,$C_s=N'/(\sqrt{\pi}\sqrt{Dt})=1.76\times10^{20}\text{cm}^{-3}$。从图1.15或者表4.1中可查出 1.5Ω·cm 衬底的掺杂浓度 $N_a=10^{16}\text{cm}^{-3}$。所以比值 $N_a/C_s=5.7\times10^{-5}$,二极管可以被看做是单边突变结。

从图2.21,根据 $C/C_s=N_a/C_s$ 得到 x/L,可求出结深 x_j。在图2.21中使用 Gauss 分布的结论,得到 $x_j=3.1\times2\sqrt{Dt}=1.0\mu\text{m}$。

(a)从图4.15或者表4.1,得到平面结的击穿电压 $BV=63\text{V}$。

(b)从图4.15,单扩散二极管击穿电压 $BV=26\text{V}$。

这个例题表明扩散结拐角电场的集中具有非常重要的实际意义。从图4.15看到随着结深的增加,击穿电压快速增大。

因为载流子在两次碰撞间连续不断地获得能量,电离率的精确表达式必须考虑两次碰撞间电场变化导致的载流子能量随位置变化而形成的能量传递的"非本地"效应。

在高性能集成电路器件中,电场随位置的变化可能非常剧烈,峰值电场只分布在几十纳米的范围内。这么剧烈的电场变化将导致碰撞电离模型仅取决于局部电场,这样将高估碰撞电离的数量。当电场随位置快速增加时,电子在低电场区只能获得部分能量,与高电场区获得的能量相比要小得多(可以说电子能量滞后于电场)。当高场区非常窄的时候,电子在这一高场区中并不能得到足够的能量来产生碰撞电离,击穿发生在更高的电压下。这样,假设位置 x 处的电子能量和本地电场相对应的模型就会对碰撞电离所产生的电流过高估计几个数量级。这一效应在理解高性能双极晶体管击穿和估算 MOS 晶体管衬底电流时意义重大。

另一个对碰撞电离建模的方法是计算平均电子能量,确定电子从上一次碰撞到发生碰撞电离所对应的能量,而不是对应于本地电场。位置 x 处的电子能量可从电场分布的能量平衡变量求出,等于电场和指数衰减长度 λ_e[8]的卷积

$$E = \frac{2q}{5}\int_0^x \mathscr{E}(u)\exp\left(\frac{u-x}{\lambda_e}\right)\mathrm{d}u \tag{4.4.17}$$

式中 λ_e 是能量与电场平衡所要求的与距离相关的能量弛豫长度,其典型值约为 60 ~ 80nm,对应的时间不到一个皮秒。电离率可以表示为这一平均能量的函数,而不再是局部电场的函数(如等式(4.4.13))。

4.4.2 Zener 击穿†

当掺杂浓度增加时,耗尽区宽度减小,发生雪崩击穿的临界电场也相应增加。在高掺杂时,产生雪崩击穿所需要的电场高于 Zener 击穿所需要的电场,使得后者发生的可能性增加。正如前面讨论的,当电场的作用力强到足以将电子从共价键中拉出来直接产生电子-空穴对时,就会发生 Zener 击穿。图 4.16 为发生 Zener 击穿时的能带图⑤,正如图中所示的,结的 p 型区价带中大量的电子,被窄的耗尽区阻隔,不能去填充 n 型区导带中具有相同能量的空的允许态能级。当半导体中的掺杂浓度增加时,给定反偏电压下的耗尽区宽度减小,耗尽区的能带弯曲更加陡峭。因为电子的波动特性,处于 p 型半导体价带中的电子有一定的概率可以隧穿通过禁带,到达 n 型半导体导带中具有相同能量的能级上。因为电子穿过势垒的透射概率强烈地依赖于势垒宽度,所以隧穿仅在高掺杂材料中才显得重要,这时电场足够高,耗尽区宽度也足够窄。

为了研究从价带到导带的隧穿几率,我们将电子遇到的势垒近似为三角形,如图 4.17 所示。势垒的高度 E_B 从 $x=0$ 处的 E_g 开始线性递减到 $x=L$ 处的 0,平均电场 $\mathscr{E}=E_g/qL$。利用等式中的势垒高度,隧穿几率 Θ 可以近似为⑥

$$\Theta \approx \exp\left[-2\int_0^L \sqrt{\frac{2m^* E_B}{\hbar^2}}\,\mathrm{d}x\right] \tag{4.4.18}$$

积分后得到的隧穿几率为

---

⑤ 在这一讨论中,需要同时考虑导带和价带中的电子,而不能认为价带中只有空穴才具有重要的地位。

⑥ 这一"WKB"近似在大部分的量子力学教科书上都会论及,例如文献[9]。

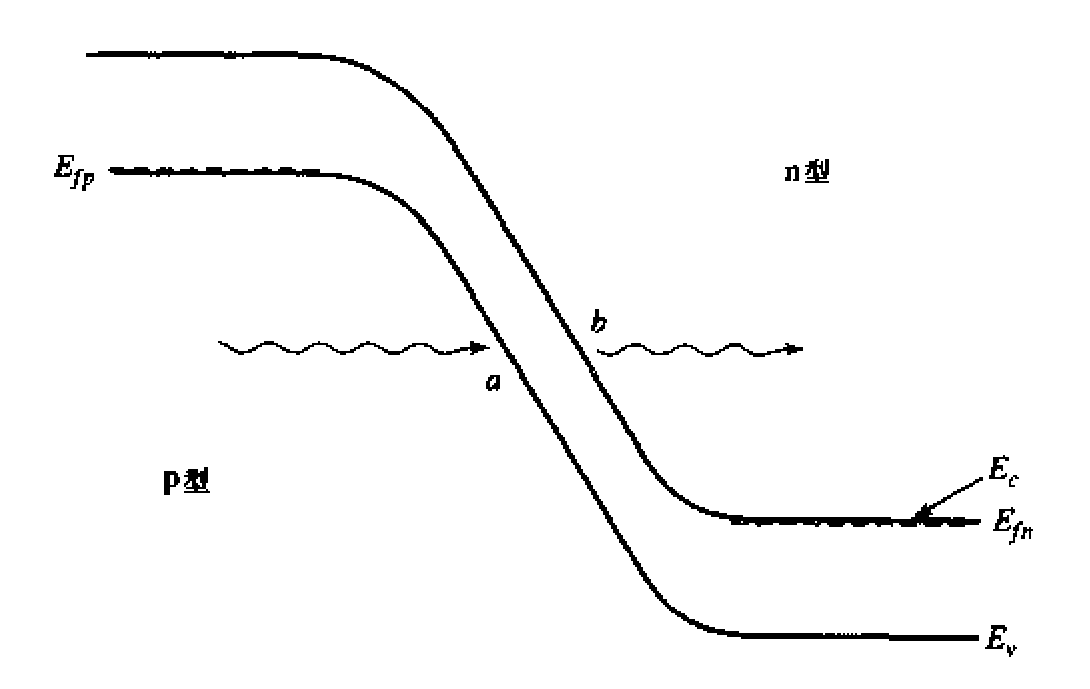


图4.16 两侧高掺杂的反偏结的能带图：隧道或Zener击穿容易发生在这种结中

$$\Theta = \exp\left(-\frac{B}{\mathscr{E}}\right) = \exp\left(-\frac{qBL}{E_g}\right) \tag{4.4.19}$$

其中

$$B = \frac{4\sqrt{2m^*}E_g^{3/2}}{3q\hbar} \tag{4.4.20}$$

隧穿几率随电场的减小，或者说随隧穿距离的增加，而快速减小。

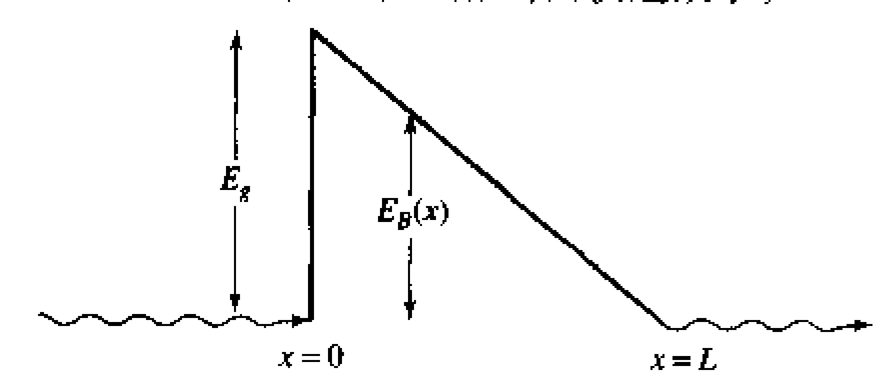


图4.17 结的隧穿几率可以近似为隧穿通过三角形势垒的几率

从上述近似可粗略估算出达到一定的隧穿所需要的电场。电流等于面积、电子电荷、p型区价带中每秒达到势垒且势垒对面存在空状态的电子数以及每个电子隧穿通过势垒几率的乘积。电子到达势垒的数量可以表示成价带电子浓度 $\mathscr{N}$ 乘以它们的速度 v，于是电流可以表示为

$$I = qA\mathscr{N}v\Theta \tag{4.4.21}$$

我们将 10mA 的电流流过 10^{-5}cm^2 面积的结作为可观隧穿电流的标准。与势垒另一侧导带中的空态能量相当的价带中的电子浓度与原子密度可比拟，大约为 10^{22}cm^{-3}。假设电子的热运动速度大约为 10^7cm s^{-1}，这样就有 $10^{29}\text{cm}^{-2}\text{s}^{-1}$ 的电子试图越过势垒。对应于这一电流的隧穿几率可以从等式(4.4.19)中得到，大约是 10^{-7}。利用这些数值，并取半导体带隙为 1eV，得到相应的隧穿距离和电场大约分别为 4nm 和 10^6V cm^{-1}。也就是说，隧道或 Zener 击穿要产生可观的漂移电流，有效的势垒必须小于 4nm，且耗尽区电场必须大于 10^6V cm^{-1}，对应的 *pn* 结轻掺杂一侧的浓度要求大于或等于 10^{18}cm^{-3}。这些数值和实测值是一致的。

当掺杂浓度减小时，空间电荷区宽度增加，隧穿几率迅速减小，雪崩击穿的发生几率大于Zener击穿。因此，Zener击穿仅发生在重掺杂结，轻掺杂结只发生雪崩击穿。

产生Zener击穿的器件，其击穿电压通常比雪崩击穿电压低。硅二极管中发生纯的Zener击穿的$BV<5V$。电压更高时，雪崩击穿通常占据主导地位。具有明确击穿特性的商用二极管无论击穿机理如何，通常都称做Zener二极管。

要确定结中到底发生的是雪崩击穿还是Zener击穿，只要观察击穿电压的温度系数即可。虽然不是很明显，但是这两种击穿电压随温度的变化趋势截然相反。Zener击穿电压随温度增加而减小，因为可隧穿的价带电子流随温度的增加而增加。而雪崩击穿的温度效应正好相反，击穿电压随温度的增加而增加，因为高能电子的平均自由程(等式(4.4.10)中的λ)减小了。击穿电压大约在5V到6V的范围内，雪崩和Zener击穿可以同时发生，这样结的击穿电压随温度的变化非常小。这一特性对设计集成电路中的参考电压源非常有用。

4.5 器件：结型场效应晶体管

4.5.1 pn结场效应晶体管(JFET)

我们已经知道在结上加反偏电压可以改变pn结中耗尽层宽度。本节将研究利用这一机理控制流过两个pn结之间区域的电流的器件。因为只有少量的电流流过反偏pn结，所以控制电极上的功耗非常小，而更多的功率通过控制电流得以传递。这种结型场效应晶体管(JFET)可用做功率放大器。

考虑图4.18所示的结构，p型衬底上有一层轻掺杂的n型层。由2.6节知，p型衬底上外延生长一层n型材料可得到这种结构，这时整个n型外延层均匀掺杂。另一种获得n型区的方法是采用离子注入工艺。

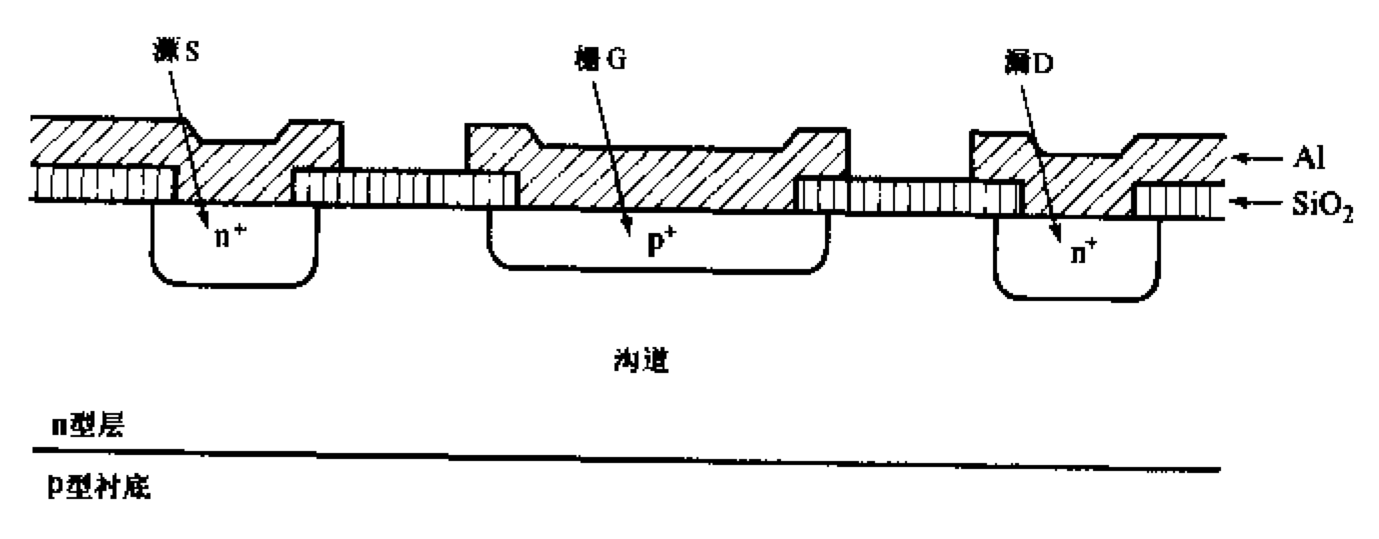


图4.18 n沟道结型场效应晶体管的基本结构。轻掺杂的n型沟道位于p型衬底和扩散形成的p型栅极之间

制备好均匀轻掺杂的n型层后，通过扩散形成两个重掺杂的n型区(记为n^+)，这样就可以在称为沟道的轻掺杂区上形成良好的欧姆接触(见3.4节)。注意沟道区与2.10节中介绍的电阻结构很相似，只是还存在p^+层。n^+电极称为源极和漏极。源极向沟道提供多数载流子，所以在n沟道JFET中，通常电流从漏端流向源端。如果改变图4.18中的所有掺杂类型，这个图就变成了p沟道JFET。通常p沟道器件的电流是从源端流向漏端。

当反偏电压加在被称做栅的 p 型区时，图 4.18 中沟道上方的 pn 结将起到控制作用。沟道定义为栅极的耗尽区和衬底 pn 结耗尽区之间的中性区。衬底通常接地。如果 n 沟道 JFET 的漏极正偏，电流便通过沟道从漏极流向源极。如果将源极接地，同时在 p^+ 极上加负偏压，pn 结的耗尽区将扩大，沟道变窄，沟道电阻增加，导致从漏端流向源端的电流减小。通过这种方式，栅极上外加电压可以控制沟道中的电流。

通过对 JFET 的工作原理的定性解释，可以定性讨论这一类器件，得到的理论将可直接应用于 GaAs 金属-半导体场效应晶体管(MESFET)，而且对今后讨论金属-氧化物-半导体场效应晶体管(MOSFET)也有益处(第 9 和第 10 章)。

JFET 器件分析 为了分析 JFET，我们先考虑源极接地，漏极上加小的偏压 V_D，栅极电压为 V_G。漏极电压和栅-沟道偏压很小，所以栅耗尽区宽度在整个沟道上是均匀的。图 4.19 是沟道区的放大图。考虑一维结构，假设源区和漏区之间的栅长为 L，垂直于纸面的沟道宽度为 W(通常 $W \gg L$)。漏极电流沿 y 方向流动。考虑 p 型栅区掺杂浓度 N_a 远大于沟道区掺杂 N_d 的单边突变结，耗尽区主要扩展在 n 型沟道区。设 p 型栅极和衬底之间的距离为 t，n 沟道区中的栅耗尽区宽度为 x_d，沟道中中性区厚度为 x_w。为重点讨论栅的影响，假设衬底结的耗尽区主要向衬底延伸，这样 $x_w \approx (t - x_d)$。这一假设在衬底轻掺杂时近似有效。

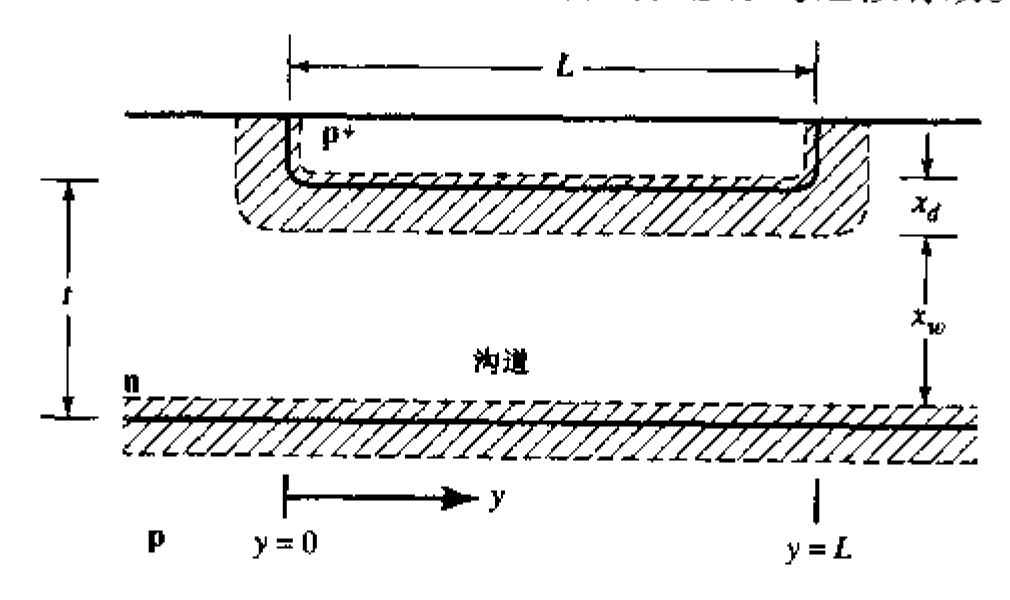


图 4.19 栅长为 L 的 JFET 的沟道区。耗尽区用阴影表示

沟道电阻可以写做

$$R = \frac{\rho L}{x_w W} \tag{4.5.1}$$

其中 $\rho = (q\mu N_d)^{-1}$ 是沟道的电阻率。于是漏极电流等于

$$I_D = \frac{V_D}{R} = \left(\frac{W}{L}\right)(q\mu_n N_d x_w V_D) \tag{4.5.2}$$

等式(4.5.2)中与栅极电压的依赖关系可由 $x_w = t - x_d$ 得到，其中 x_d(由等式(4.3.1))为

$$x_d = \left[\frac{2\epsilon_s}{qN_d}(\phi_i - V_G)\right]^{1/2} \tag{4.5.3}$$

式中 ϕ_i 是内建电势。这样电流就可以写成栅极电压和漏极电压的函数

$$I_D = \frac{W}{L} q\mu_n N_d t \left\{1 - \left[\frac{2\epsilon_s}{qN_d t^2}(\phi_i - V_G)\right]^{1/2}\right\} V_D \tag{4.5.4}$$

括号前的因子等于 n 型区完全没有耗尽区时(所谓的冶金沟道)的电导 G_0。等式(4.5.4)可以重新表示为

$$I_D = G_0\left\{1 - \left[\frac{2\epsilon_s}{qN_d t^2}(\phi_i - V_G)\right]^{1/2}\right\} V_D \qquad (4.5.5)$$

在给定栅极电压下,并且外加漏极电压很小时,可得到 I_D 和 V_D 之间呈线性关系。而根据栅-沟道结为突变结的条件,等式(4.5.5)中的漏极电流与栅极电压成平方根关系。等式(4.5.5)还表明电流在外加栅极电压等于零时最大,随 $|V_G|$ 的增加而减小。并且当栅极电压大到足以使整个沟道区耗尽时,电流将变为零。

例题　JFET 中的源和漏电阻

从图 4.18 的 JFET 横截面图可看到,重掺杂的源电极和漏电极与沟道区是分离的,这一分离增大了源漏电极与沟道的串联电阻。考虑 n 沟道 JFET,$L=5\mu m$,$W=10\mu m$,沟道区和源漏扩散区之间的距离 L' 是 $5\mu m$。已知 n 型外延层掺杂浓度 $N_d=5\times10^{15}cm^{-3}$,外延层厚度 $t_e=1.5\mu m$。p 型栅区掺杂浓度 $N_a=1\times10^{19}cm^{-3}$,厚度为 $0.5\mu m$。假设栅扩散构成突变结,同时忽略衬底结在沟道区的延伸,令 $V_G=0$。

求由源极和漏极的串联电阻(考虑线性区工作的 JFET)使沟道电阻增加的百分比。

解:沟道电导由等式(4.5.4)给出。利用等式(4.2.10)求出 $\phi_i=0.854V$。根据表 4.1 或者图 1.15,查出沟道区电阻率 $\rho_n=1\Omega\cdot cm$,其中 $\rho_n=(q\mu N_d)^{-1}$。将这些值代入等式(4.5.4)中,并考虑 $W/L=2$,$t=1\mu m$,可计算出沟道电导 G_c

$$G_c = (2.08\times10^{-4})(1-0.472) = 1.1\times10^{-4}\Omega^{-1} = 0.11mS$$

其中电导单位是西门子。

连接源极(或漏极)到沟道区的串联电阻的电导 $G_{s,d}$ 为

$$G_{s,d} = (W/L)q\mu N_d t_e = 3.1\times10^{-4}\Omega^{-1} = 0.31mS$$

沟道电阻是 $1/G_c=9.1k\Omega$,源极和漏极的串联电阻分别是 $1/G_{s,d}=3.2k\Omega$,所以总电阻就是 $15.5k\Omega$,增加的百分比为 $6.4/9.1\times100=70\%$。

增加的电阻是寄生电阻,它们与输出电压串连,减少了 JFET 的增益。本题条件下(外加栅极电压为零时电流最大)的寄生电阻效应最大。当栅极加反向偏压时,G_c 减小,导致寄生电阻效应减小。

现在考虑任意 V_D 和 V_G(栅必须保持反偏)下的其他物理效应。考虑 V_D 对沟道电位的影响,那么沟道和栅之间的电压是 y 的函数,因此,耗尽区宽度和沟道横截面积都随 y 变化。耗尽区靠近漏端的电压比源端高,因此耗尽区宽度在漏端要宽一些,如图 4.20 所示。

采用缓变沟道近似,即假设从源端到漏端,沟道和耗尽层宽度变化缓慢,这样耗尽区仅受垂直方向电场的影响,而不受从源到漏的横向电场的干扰。换句话说,就是耗尽区电场在 y 方向的分量比 x 方向的分量小很多。在这一近似下,可以从一维分析得到耗尽区宽度。

y 处沟道长度 dy 上的电压增量为

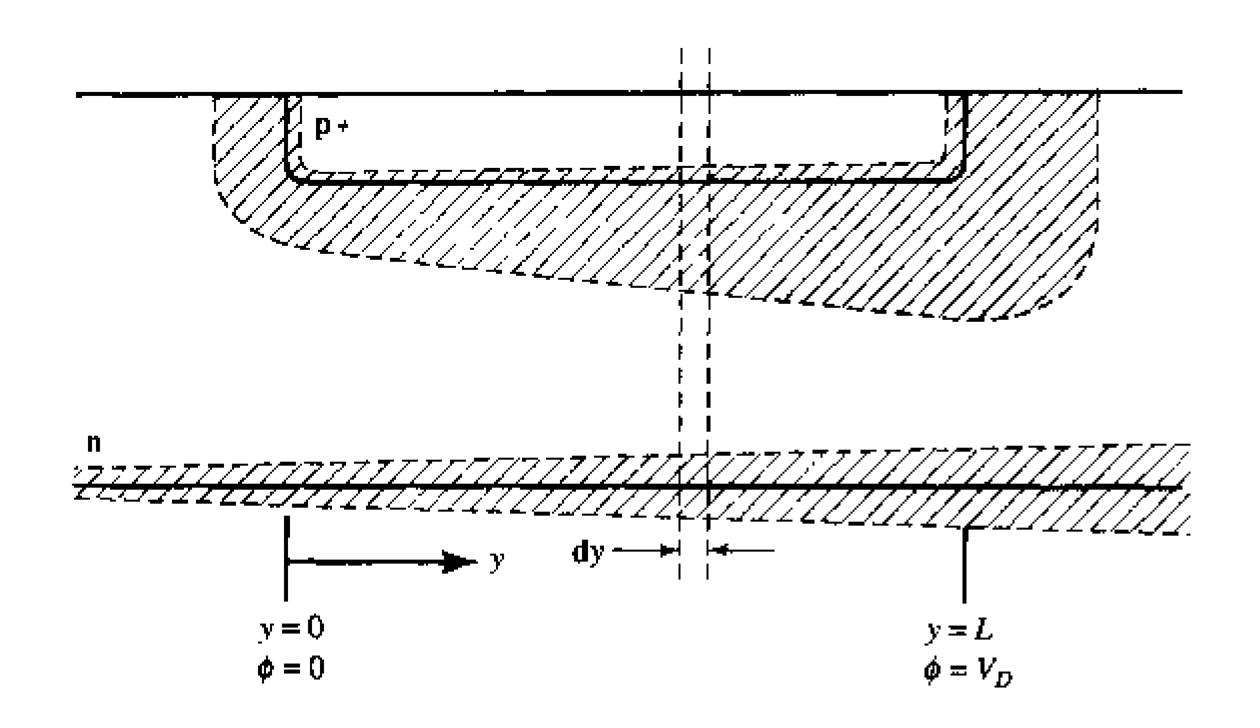


图4.20 当漏端电压相对源端较高时,JFET沟道中耗尽区宽度沿沟道方向发生变化

$$d\phi = I_D dR = \frac{I_D dy}{Wq\mu_n N_d(t - x_d)} \tag{4.5.6}$$

耗尽区宽度 x_d 受到电压 $\phi_i - V_G + \phi(y)$ 的控制,这里 $\phi(y)$ 是 y 处的沟道电势,这样

$$x_d = \left[\frac{2\epsilon_s}{qN_d}(\phi_i - V_G + \phi(y))\right]^{1/2} \tag{4.5.7}$$

将等式(4.5.7)代入等式(4.5.6),并从源端到漏端积分,得到JFET的电流电压关系

$$\frac{I_D \int_0^L dy}{Wq\mu_n N_d} = \int_0^{V_D} \left\{t - \left[\frac{2\epsilon_s}{qN_d}(\phi_i - V_G + \phi)\right]^{1/2}\right\} d\phi \tag{4.5.8}$$

求出积分并重新整理得到

$$I_D = G_0\left\{V_D - \frac{2}{3}\left(\frac{2\epsilon_s}{qN_d t^2}\right)^{1/2}\left[(\phi_i - V_G + V_D)^{3/2} - (\phi_i - V_G)^{3/2}\right]\right\} \tag{4.5.9}$$

在漏极电压较小的情况下,等式(4.5.9)可简化成等式(4.5.5),电流随漏极电压线性增加;然而,当漏极电压较大时,电流随漏极电压的增加变缓。

在漏极电压较大时,等式(4.5.9)表明电流将达到最大值,然后开始随着漏极电压增加而减小,但这是不符合实际情况的。从图4.21可看出,当漏极电压增加时,漏端导电沟道的宽度减小,直至该区域的沟道完全耗尽,如图4.21(b)所示。当这一现象发生时,等式(4.5.6)不再成立($x_d \to t$),所以这个等式仅在 V_D 小于沟道夹断时的漏极电压时才有效。当沟道被夹断时,仍然有电流流过,因为电子通过沟道向漏端运动时没有势垒阻挡,当它们到达夹断区边界(夹断点)时,被从漏端指向源端的电场直接拉到漏端。如果漏端电压继续增加,更多的电压将会降落在漏端附近的这一耗尽的高场区,沟道中完全耗尽点的位置将向源端移动,如图4.21(c)所示。如果这一小的移动可被忽略,那么当漏极电压增加时,漏极电流将保持不变(饱和),这种现象被称为饱和。漏端附近沟道完全耗尽时对应的漏极电压(漏极饱和电压 V_{Dsat})可从等式(4.5.7)得到

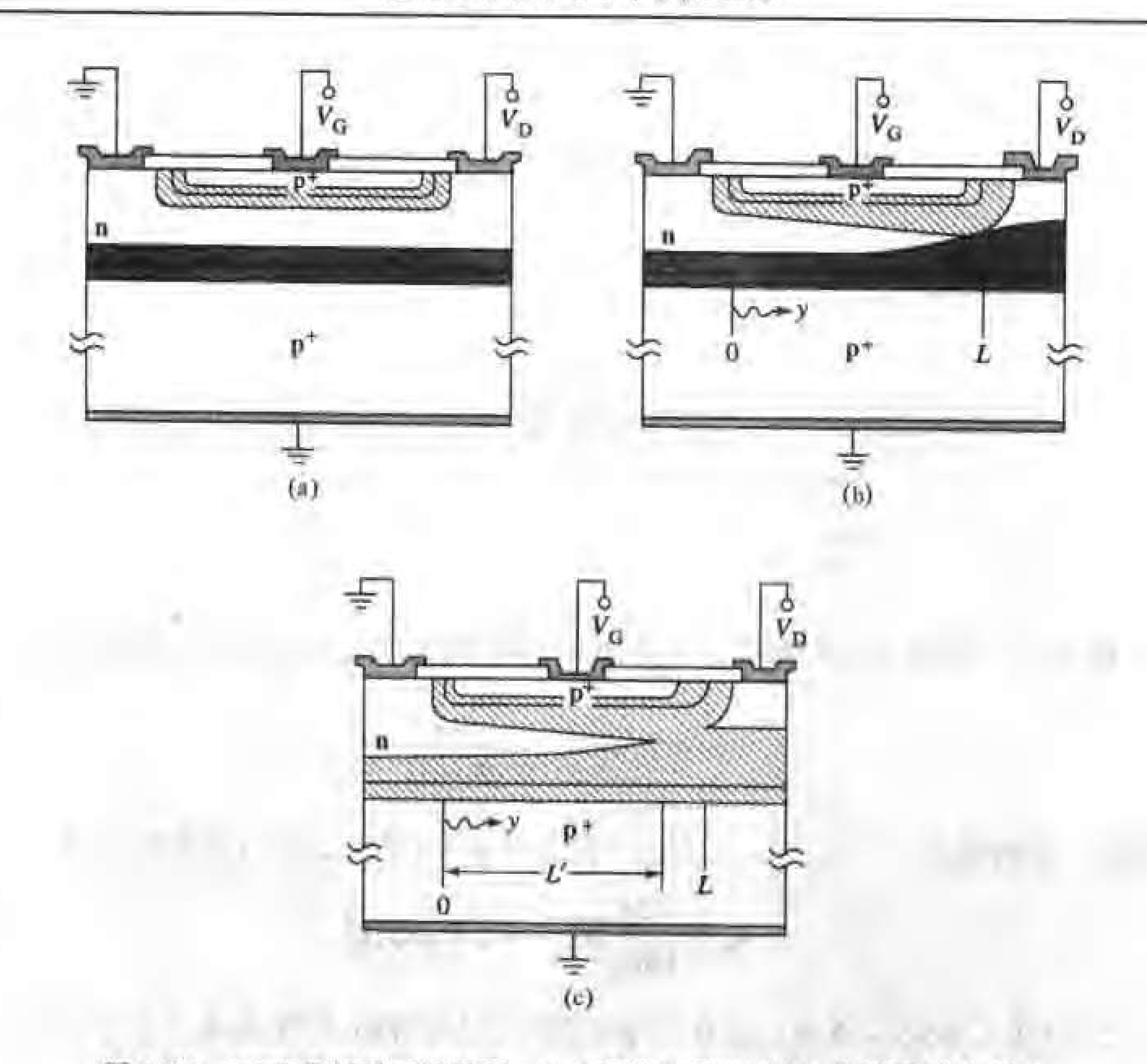


图4.21 JFET中耗尽区的性质。(a)漏极电压较小时,沟道区近似等电位,耗尽区宽度不变;(b)当 V_D 增加到 V_{Dsat} 时,沟道两侧的耗尽区在夹断点相遇($y=L$ 处);(c)当 $V_D>V_{Dsat}$ 时,夹断点($y=L'$处)向源端移动[10]

$$V_{Dsat}=\frac{qN_dt^2}{2\epsilon_s}-(\phi_i-V_G) \tag{4.5.10}$$

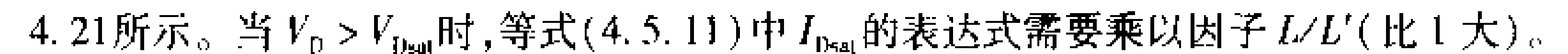
与之对应的饱和漏电流是

$$I_{Dsat}=G_0\left[\frac{qN_dt^2}{6\epsilon_s}-(\phi_i-V_G)\left\{1-\frac{2}{3}\left[\frac{2\epsilon_s(\phi_i-V_G)}{qN_dt^2}\right]^{1/2}\right\}\right] \tag{4.5.11}$$

在上述分析的基础上,可将漏极电流-漏极电压特性划分成三个区域(图4.22):(1)低漏极电压对应的线性区,(2)电流随漏极电压增加的速率低于线性区的区域,(3)漏极电压进一步增加时电流保持为常数的饱和区。与物理分析一样,等式(4.5.11)表明,饱和漏极电流在栅偏压为零时最大,且随外加负栅电压的增大而减小,相应的饱和漏极电压也减小。图4.22为不同栅压下的JFET的输出特性曲线。负栅电压足够大时,饱和漏极电流将减小至零。这一关断电压 V_T 可以从等式(4.5.11)得到

$$V_T=\phi_i-\frac{qN_dt^2}{2\epsilon_s} \tag{4.5.12}$$

实际上当漏极电压 $V_D>V_{Dsat}$ 时,漏极电流随漏极电压的增加而略微增加,这是因为等式(4.5.8)的积分上限变为 L',而不是 L。这里 L' 是沟道完全耗尽处对应的位置[$\phi(L')=V_{Dsat}$],如图



4.21所示。当 $V_D > V_{Dsat}$时,等式(4.5.11)中 I_{Dsat}的表达式需要乘以因子 L/L'(比 1 大)。

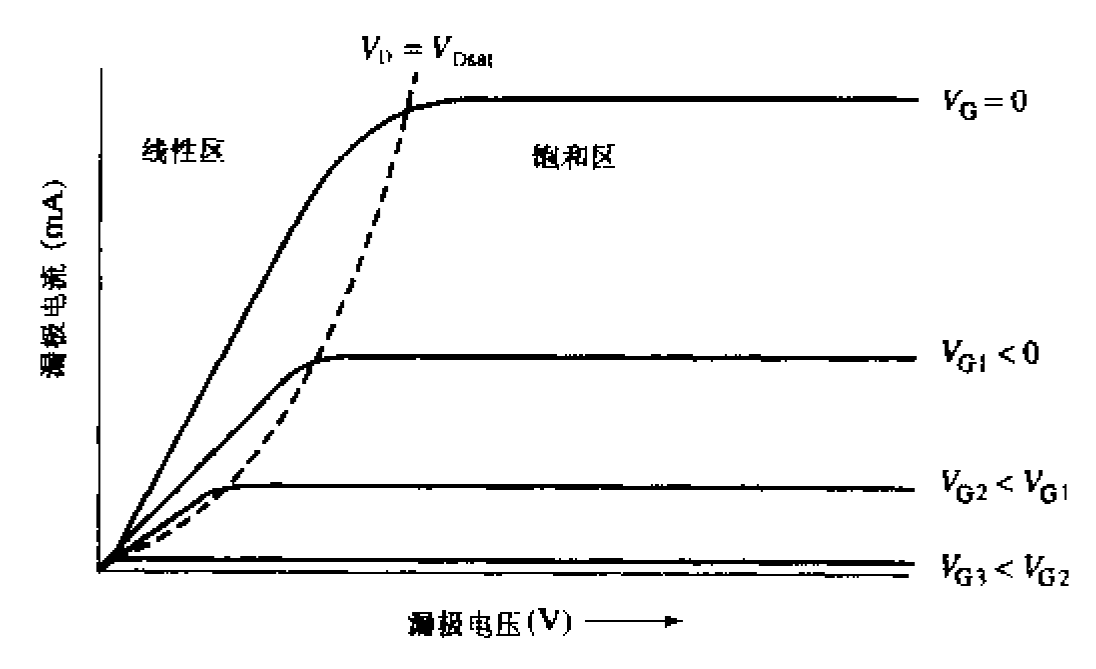


图 4.22 JFET 在不同栅极电压下的输出特性(漏极电流-漏极电压)曲线。线性区(曲线 $V_D = V_{Dsat}$ 曲线左侧)对应于图 4.21a,饱和区($V_D = V_{Dsat}$曲线右侧)对应于图 4.21(c)

尽管大部分 JFET 表现出如图 4.22 所示的较好的饱和区特性,但是短沟道器件的饱和特性却有较大的偏离。沟道施主浓度为 10^{16}cm^{-3}时,比 V_{Dsat}高 5V 的电压将使沟道耗尽区增加 1.0μm。当沟道长度为 8μm 时,比值 $L/L' \approx 8/7$,与前面的结果偏差不大;当沟道长度为 2μm 时,就会发生较大的偏差[①]。

场效应晶体管通常工作在饱和区,输出电流不会受到输出(漏极)电压的显著影响,输入(栅极)电压起主要控制作用。在这一偏置条件下,JFET 基本上是输入电压控制的理想电流源。晶体管的跨导 g_m反映了栅极电压对漏极电流的控制。其定义为

$$g_m = \left.\frac{\partial I_D}{\partial V_G}\right|_{V_D = \text{const}} \tag{4.5.13}$$

对等式(4.5.9)微分得到 g_m 表达式

$$g_m = G_0\left(\frac{2\epsilon_s}{qN_d t^2}\right)^{1/2}\left[(\phi_i - V_G + V_D)^{1/2} - (\phi_i - V_G)^{1/2}\right] \tag{4.5.14}$$

g_m在饱和区达到最大值,由等式(4.5.14)得

$$g_{m\text{sat}} = G_0\left[1 - \left(\frac{2\epsilon_s}{qN_d t^2}(\phi_i - V_G)\right)^{1/2}\right] \tag{4.5.15}$$

上述分析包括了几个假设,但在实际器件中,部分假设可能不成立,使得理论和实验不能很好地吻合。例如,假设耗尽层宽度由栅-沟道结控制,而不受衬底-沟道结影响。然而,衬底-沟道结上的电势沿沟道方向也变化,在靠近漏端处电势最大,耗尽层最宽。因此,在漏极电压比等式(4.5.10)计算得到的电压低时沟道就会完全耗尽。加在衬底和源极之间的电压有时候也被用来控制 JFET 的特性。在等式(4.5.6)中加入衬底偏置对未耗尽的沟道厚度的调制作用,可以很容易计算出衬偏效应。

---

① 在第 9 章介绍金属-氧化物-半导体场效应晶体管时将进一步讨论沟道长度调制效应和短沟道效应。

考虑有源沟道和重掺杂的源漏电极之间的轻掺杂区,会让问题变得棘手。因为光刻的限制和击穿电压的需求,通常 n^+ 扩散区和 p 型栅扩散区在空间是分开的。正如前面的例题看到的,源漏串联电阻的存在将导致器件特性偏离理想特性,尤其是在电流较高时,因此在分析和设计实际器件时必须要加以考虑。

我们已经看到,JFET 的特性对沟道区的厚度 t 和掺杂浓度极度敏感。n 型沟道由可精确控制的外延生长形成。更为关键的工艺是 p 型栅区的扩散。这一扩散会使有效沟道厚度不易控制。为了提高控制精度,通常采用离子注入引进栅杂质。如果集成电路中的其他器件不需要外延工艺,JFET 也可以通过两次注入来形成,一次是 n 沟道掺杂,一次是 p 型栅区掺杂。

图 4.23 是包含几个 JFET 的集成电路版图(355 运算放大器)。版图下方的交叉梳状结构是大输入 JFET。为了降低源漏电阻(大的 W/L),源和漏制成梳状结构。栅电极分布在梳状结构的源和漏之间。

4.5.2 金属-半导体场效应晶体管(MESFET)

在硅集成电路中,通过离子注入或者外延生长可以很容易制造出 pn 结。前面看到 pn 结耗尽区的厚度可以调制导电区的厚度,从而控制流过沟道的电流,形成有用的晶体管。

而在其他材料体系中,形成 pn 结并不总是那么容易。这时,可利用 Schottky 势垒,而不是 pn 结来形成随电压变化的耗尽区,如图 4.24 所示。这种器件称为金属-半导体场效应晶体管(MESFET),利用加在 Schottky 势垒上的电压控制沟道中电流的大小。

Schottky 势垒可以看做是单边 pn 结的极端情况,这样大部分关于 JFET 的分析就可以应用到 MESFET 上,其中最主要的修正是内建电势 ϕ_i。图 3.4 中所示的理想金属-半导体接触的内建电势为

$$\phi_i = \Phi_M - \Phi_S = \Phi_M - X - \frac{E_c - E_f}{q} \tag{4.5.16}$$

实际上的金属半导体接触的表面态将导致 ϕ_i 偏离理想值,正如 3.5 节讨论的,通常需要用经验值确定。

当电路中包含大量的 MESFET 元件时,相当可观的功耗将限制芯片上晶体管的数量。为了减少功耗,可以调整阈值电压,使栅极电压为零时没有电流流过。适当选择沟道厚度和掺杂浓度,使得当 $V_G=0$ 时沟道完全耗尽,这种器件称为增强型器件。为使 $V_G=0$ 时,$I_D=0$,从等式(4.5.5)得到

$$N_d t^2 = \frac{2\epsilon\phi_i}{q} \tag{4.5.17}$$

当外加小的正栅压时(对 n 沟道晶体管来说为正),沟道开启,产生漏极电流。正的栅偏压必须较小(尤其是要比 ϕ_i 小),以防止产生明显的栅极电流。因为这类器件所允许的栅极电压范围很窄,所以必须精确控制沟道厚度和掺杂。

在 GaAs 集成电路工艺中,MESFET 非常有用。这种材料系统形成 Schottky 势垒比制作 pn 结更容易。GaAs 中的电子迁移率很高,适于制作超高频晶体管。金属栅电极的电阻很低,有助于提高高频性能。与硅相比,GaAs 的电阻率可以很高,这样导电沟道下方的区域就能有非常高的电阻。因此,与衬底-沟道结相关的电容非常小,从而进一步提高了晶体管的频率响应。

此外，GaAs 中 ϕ_t 值比硅大，使得 GaAs 增强型 MESFET 器件的 V_G 的变化范围更宽。

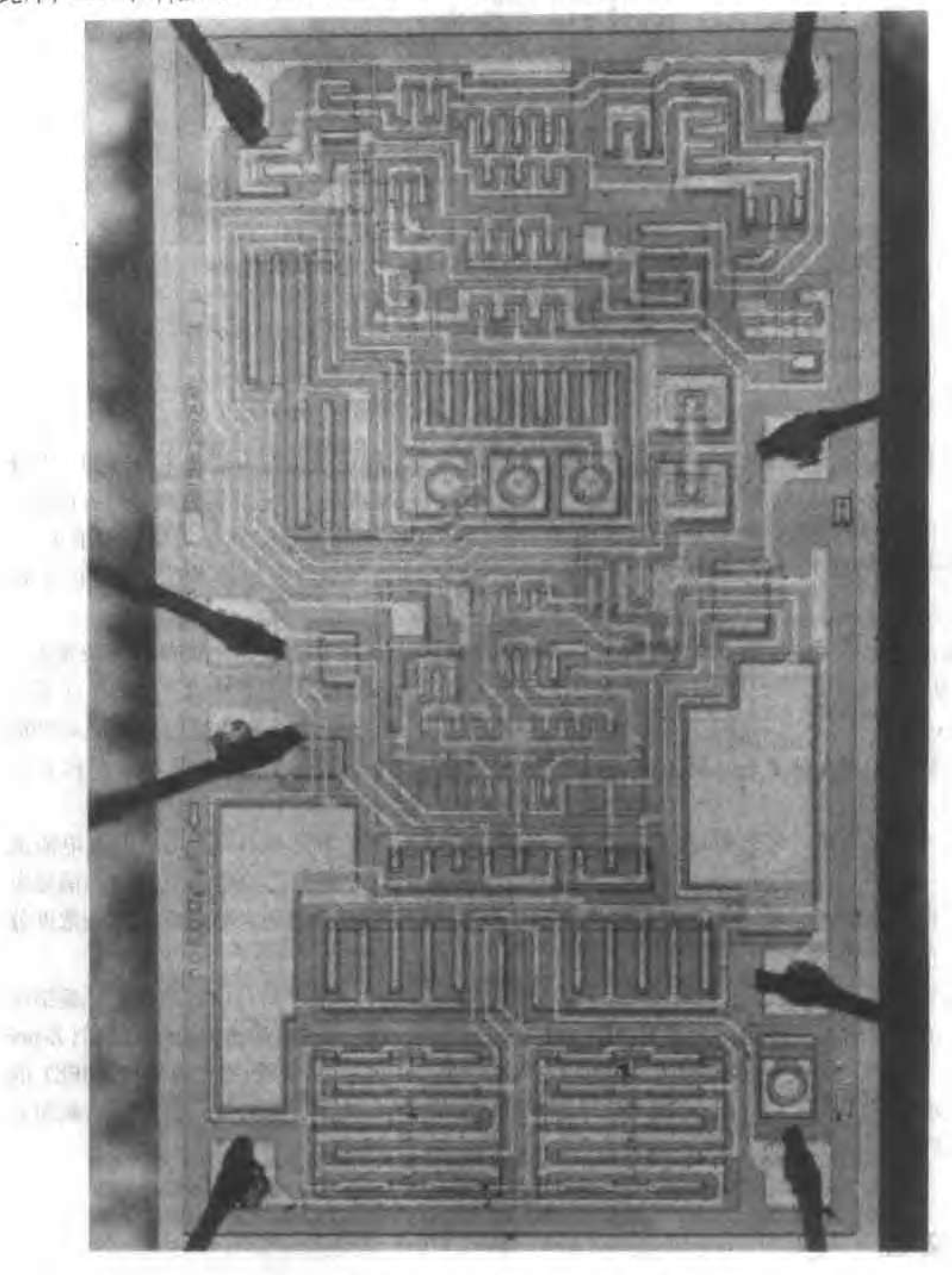


图 4.23 可获得高输入电阻的 JFET 集成电路（运算放大器）（引自 National Semicondutor Corp.）

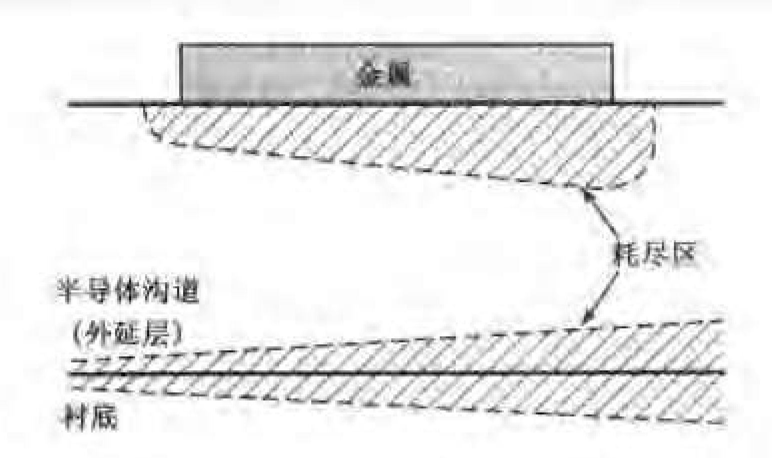


图 4.24 金属-半导体场效应晶体管(MESFET)的横截面图,其中阴影部分为金属栅下以及衬底-沟道界面处的耗尽区

小结

与金属-半导体接触的情况类似,也可以通过热平衡原理得到非均匀掺杂半导体的一些重要性质。掺杂不均匀的同一种类型的半导体,也会形成内建电场,由它产生的漂移趋势可平衡自由载流子的扩散趋势。通常,与这个电场相关的空间电荷很少,半导体可近似看做准中性,多数载流子浓度等于净掺杂浓度。当掺杂浓度梯度增大时,准中性近似将不再有效。由于 pn 结的掺杂梯度通常很大,所以不能用准中性近似。

通常用耗尽近似分析 pn 结的效应。在耗尽近似下,假定空间电荷由未被补偿的杂质离子组成,并假定其边缘突变。采用这种近似得到的电场和电势分布,只是在中性区边界一个非本征 Debye 长度内是不精确的。与金属-半导体结类似,平衡时 pn 结存在内建电势。同质结的内建电势等于相互独立的 p 型区和 n 型区的 Fermi 能级之差,异质结带边必须考虑 E_c 和 E_v 的不连续。

当外加电压使内建电势增加时,例如,p 型区接地,n 型区加正电压时,几乎没有电流流过。这一反偏电压将导致 pn 结的空间电荷区加宽。在不同的直流反偏电压下测量小信号电容时很容易观察到耗尽区宽度的变化。测量一系列的电容电压关系可获得结区的掺杂浓度分布的信息。

在高场区,半导体的电导率会突然升高,因为其内部产生了额外的自由载流子。反偏结中出现电流突然增加的现象称为击穿。导致击穿的机制有两种:(1)雪崩击穿,(2)隧道(Zener 击穿)。反偏 pn 结在集成电路中的一种应用是作为结型场效应晶体管(JFET)的栅。JFET 的工作原理直接取决于对反偏 pn 结中耗尽层宽度 x_d 的调制。如果电流在受 x_d 控制的横截面上流动,这一反偏电压就可以调制电流。

参考文献

1. Courtesy W. G. OLDHAM.
2. R. PEOPLE, *Phys. Rev. B* **32**, 1405 (1985); R. PEOPLE, IEEE J. Quantum Electronics **QE-22**, 1696 (1986).
3. A. S. GROVE, *Physics and Technology of Semiconductor Devices*, Wiley, New York, 1967, p. 193. Reprinted by permission of the publisher.
4. A. G. CHYNOWETH, W. L. FELDMAN, C. A. LEE, R. A. LOGAN G. L. PEARSON, and P. AIGRAIN, *Phys. Rev.* **118**, 425 (1960).
5. S. L. MILLER, *Phys. Rev.* **105**, 1246 (1957).
6. A. S. GROVE, *ibid.*, p. 197.

7. H. L. Armstrong, *IRE Trans Electron Devices* **ED-4**, 15 (1957). Reprinted by permission of the publisher.
8. J. W. Slotboom, G. Streutker, M. J. v. Dort, P. H. Woerlee, A. Pruijmboom, and D. J. Gravesteijn, Tech. Digest International Electron Devices Meeting, (Washington, DC, December 8-11, 1991), paper 5.4, pp. 127–130.
9. E. Merzbacher, *Quantum Mechanics*, 2nd edition, Wiley, New York, 1970.
10. A. S. Grove, *ibid.*, p. 244.

参考书

D. A. Fraser, *The Physics of Semiconductor Devices*, 2nd Edition, Oxford at the Clarendon Press, 1979.

G. W. Neudeck, *The PN Junction Diode*, Vol II Modular Series on Solid-State Devices, Addison-Wesley, Reading, Mass., 1983.

习题

4.1* 已知突变硅 pn 结的掺杂浓度 $N_a=1\times10^{15}cm^{-3}$，$N_d=2\times10^{17}cm^{-3}$

(a)计算室温下的内建电势 ϕ_i。

(b)利用耗尽近似，计算结电压 V_a 等于 0V 和 -10V 下的空间电荷区宽度以及峰值电场。

4.2 已知突变硅 pn 结，一侧为重掺杂，另一侧掺杂相对较轻。轻掺杂一侧的浓度分别为(a)$10^{15}cm^{-3}$，(b)$10^{16}cm^{-3}$，(c)$10^{17}cm^{-3}$，(d)$10^{18}cm^{-3}$。求出最大场强达到图 4.14 所示的击穿场强时掺杂浓度与耗尽区域宽度的关系，并求出与之相应的外加电压。

4.3* 已知指数掺杂分布 $N=N_0\exp(-x/\lambda)$ 的半导体，其中表面掺杂浓度为 $10^{18}cm^{-3}$，$\lambda=0.4\mu m$，计算准中性区的内建电场大小。将这一电场与突变 pn 结耗尽区的最大电场做比较，假设突变 pn 结两侧的受主和施主掺杂浓度分别为 $10^{18}cm^{-3}$ 和 $10^{15}cm^{-3}$。

4.4 在 4.2 节例题中的掺杂浓度变化区域内求电势分布的函数表达式。假设过渡区两侧均匀掺杂，并画出电势分布和能带图。

4.5 (a)已知硅 p-i-n 结构的掺杂分布如图 P4.5 所示，求出并画出内建电场和内建电势分布图。指出每个耗尽区的宽度。(符号 i 代表非常轻的掺杂或者近乎本征的区域。)

(b)比较 p-i-n 结构与 pn 结的最大电场，其中 p 型区和 n 型区的掺杂浓度与(a)相同。

(c)从物理上解释本征区发生了什么情况。(也就是说，该区域耗尽近似意味着什么？)

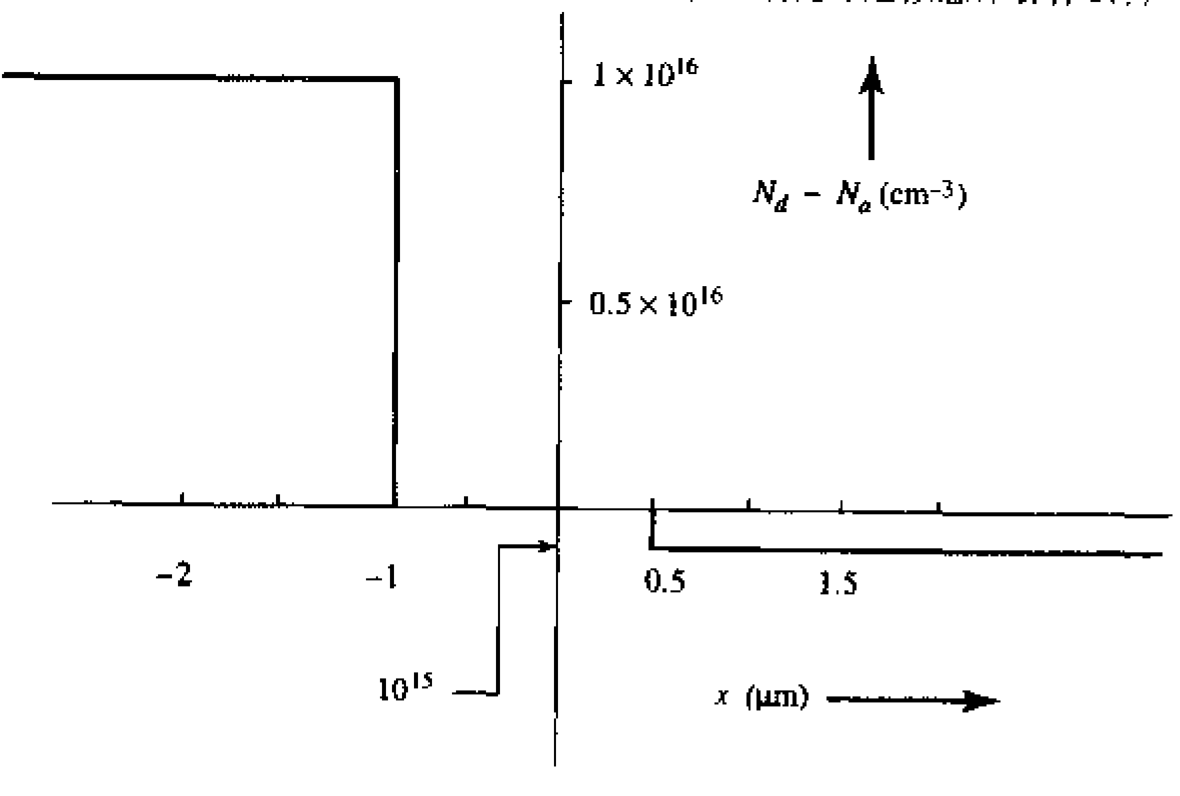


图 P4.5

(d)讨论这一结构的耗尽电容如何随电压变化,并与没有本征区但其他部分掺杂浓度相同的pn结的耗尽电容比较。画出两种情况下的$1/C^2$与外加反偏电压的关系。采用相同的坐标系,以便直接比较这两种情况。

4.6 利用耗尽近似研究线性缓变结,已知在整个耗尽区有$(N_d - N_a) = ax$。假设平衡状态的空间电荷区宽度为x_{d0}。利用给出的参数,推导以下几个表达式:(a)内建电势,(b)电场与外加电压及位置的函数关系,(c)耗尽电容与电压的函数关系。

4.7[1] 当$V_a \gg \phi_i$时,已知突变pn结的电容随$V_a^{-1/2}$变化,线性缓变结的电容随$V_a^{-1/3}$变化,这里ϕ_i是内建电势,V_a是加在结上的反偏电压。在TV调谐电路中,要求$V_a \gg \phi_i$时,需要电容随V_a^{-1}变化,定性讨论需要的掺杂分布形式。指出上述三种情况下耗尽区宽度如何随电压变化。

4.8[1] 已知硅中的掺杂分布如图P4.8所示。

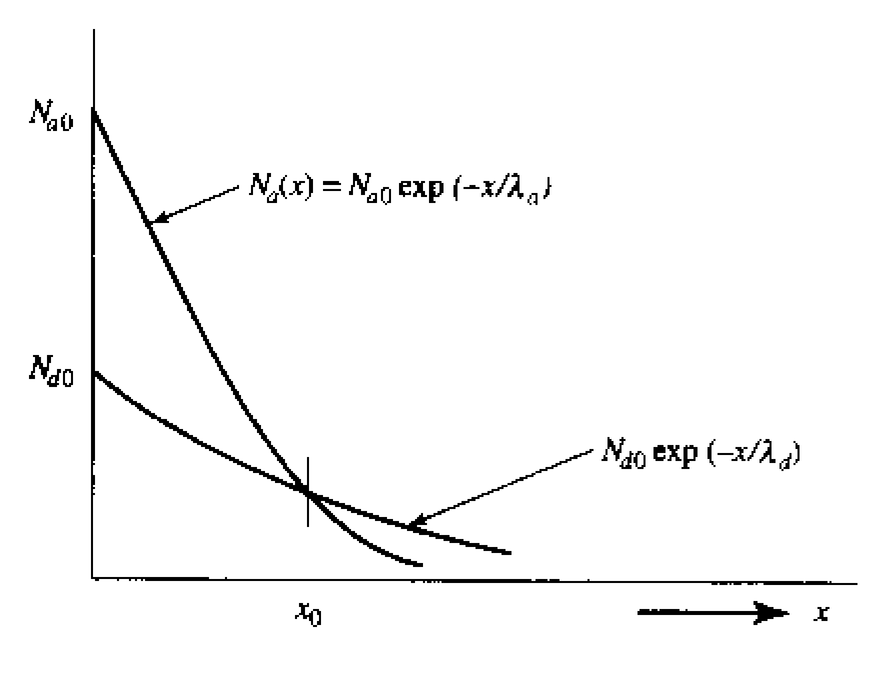


图 P4.8

(a)如果期望冶金结的位置(净掺杂浓度为零处)在$x_0 = 1\,\mu\text{m}$的地方,那么表面掺杂密度N_{a0}的值应该是多少?

(b)利用耗尽近似,画出结附近的空间电荷分布图。假设该pn结可近似为线性缓变结。选择掺杂梯度a值。

(c)在近似条件(b)下,取$\phi_i = 0.7\text{V}$,利用等式(4.3.2)和(4.3.4),计算热平衡条件下的$\mathscr{E}_{max}$,并绘出整个区域在热平衡条件下的电场(取$N_{d0} = 10^{18}\text{cm}^{-3}$, $x_0 = 10^{-4}\text{cm}$, $\lambda_a = 10^{-4}\text{cm}$, $\lambda_d = 2\times 10^{-4}\text{cm}$)

4.9 测量面积为10^{-5}cm^2的pn结小信号电容,得到的$(1/C_d^2)$与外加电压V_a的关系图如P4.9所示。

(a)如果二极管是单边突变结,求出低电导率一侧的掺杂浓度(利用曲线的斜率)。

(b)画出低电导率一侧的掺杂浓度分布图。并求出掺杂浓度改变点的位置。

(c)利用图中$(1/C_d^2)$上的截距求出高掺杂一侧的掺杂浓度。

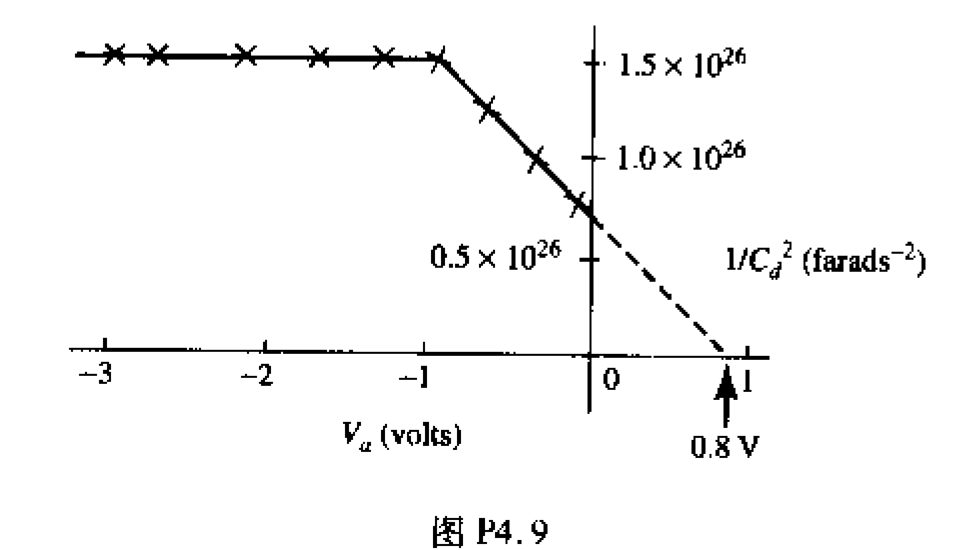


图 P4.9

4.10* 下述所有问题与图 P4.10 所示的系统有关。该图为采用硅平面工艺制造的芯片的横截面图。已知 p 型衬底的电阻率为 10Ω-cm，外延层施主掺杂浓度为 $5 \times 10^{15} cm^{-3}$，厚度为 2.5μm。A 处的铂(Pt)电极直接制备在外延硅表面。B 处 Al 淀积制备在 n^+ 层上，n^+ 区的扩散深度为 1.5μm，施主掺杂浓度为 $3 \times 10^{18} cm^{-3}$。下面的分析忽略边缘效应。

(a)求 Fermi 能级与本征 Fermi 能级(E_i)的间距：(i) n 型外延层，(ii) p 型衬底。

(b)已知测量得到的 Pt-Si 结的势垒高度(导带边与 Fermi 能级之间的能量差)是 0.85eV。

(i)求 Pt-Si 结的内建电势。

(ii)0.85eV 的势垒高度是否满足理想的 Schottky 理论，请给出解释。

(c)请问在达到击穿电场 $3 \times 10^5 V\ cm^{-1}$ 之前，Pt 电极下的 n^- 型层是否有可能完全耗尽(考虑电极 B 和电极 C 都接地，外加电压加在电极 A 上)？完全耗尽需要多高的电压？

(d)定性画出平衡时沿 1-1 方向(通过 Pt-Si 结进入衬底)的能带图，指出真空能级，并假设理想 Schottky 理论适用。

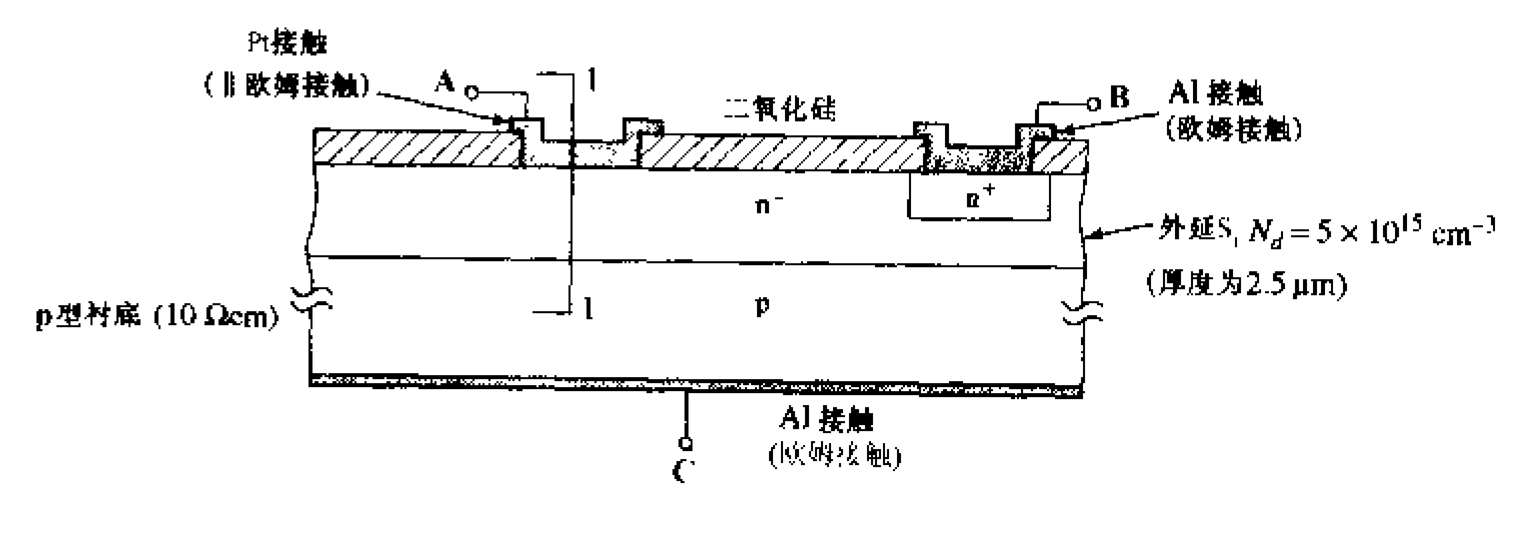


图 P4.10

4.11 用第一原理来处理雪崩，考虑入射电子碰撞晶格并释放一个电子-空穴对，假设在相互作用后，这三个粒子具有相同的动能，同时也假设它们具有相同的质量。采用能量和动量守恒原理证明雪崩发生的阈值是入射电子拥有$(3/2)E_g$的动能。(尽管考虑这个问题时采用了大量的近似，但这一能量是实际测量值的一阶近似。)

4.12 考虑反偏硅或锗 pn 结二极管，如果两者的峰值电场相同，哪种二极管更容易发生 Zener 击穿？为什么？(考虑材料的带隙。)

4.13† 根据等式(4.4.20)分析并证明导出的隧穿距离和电场值是正确的。已知电子的静止质量为 m^*，带隙为 1.1eV，等式(4.4.19)中的系数 B 为 $7.87 \times 10^7 V\ cm^{-1}$。在计算中采用电导率有效质量。

4.14 在固定栅电压下，推导 JFET 器件的线性区电导随温度变化的表达式。假设迁移率随 $T^{-3/2}$ 变化。

4.15† 推导给定栅电压下 JFET 器件饱和区漏端电导 $g = \partial I_D / \partial V_D$ 的表达式。假设这个电导缘于漏端附近耗尽区的加宽，并假定漏区重掺杂，形成的结为一维突变结。

4.16† 图 P4.16 表示一个环状结构的 JFET 器件。其中的 pn 结可以近似为线性缓变结，掺杂梯度 $dN/dx = a$，忽略源漏端串联电阻。

(a)根据上述条件求出栅关断电压 V_T？

(b)写出 I_D 关于 V_D 和 V_G 的微分关系式，并写出求解 I_D 的定积分形式。

(c)求出跨导 $\partial I / \partial V_G$ 与半径 r_1 和 r_2 的关系。已知器件的 $r_1 = 10\mu m, r_2 = 40\mu m$ 时，$g_m = 10mS$。如果其他参数不变，取 $r_2 = 60\mu m$，求 g_m。

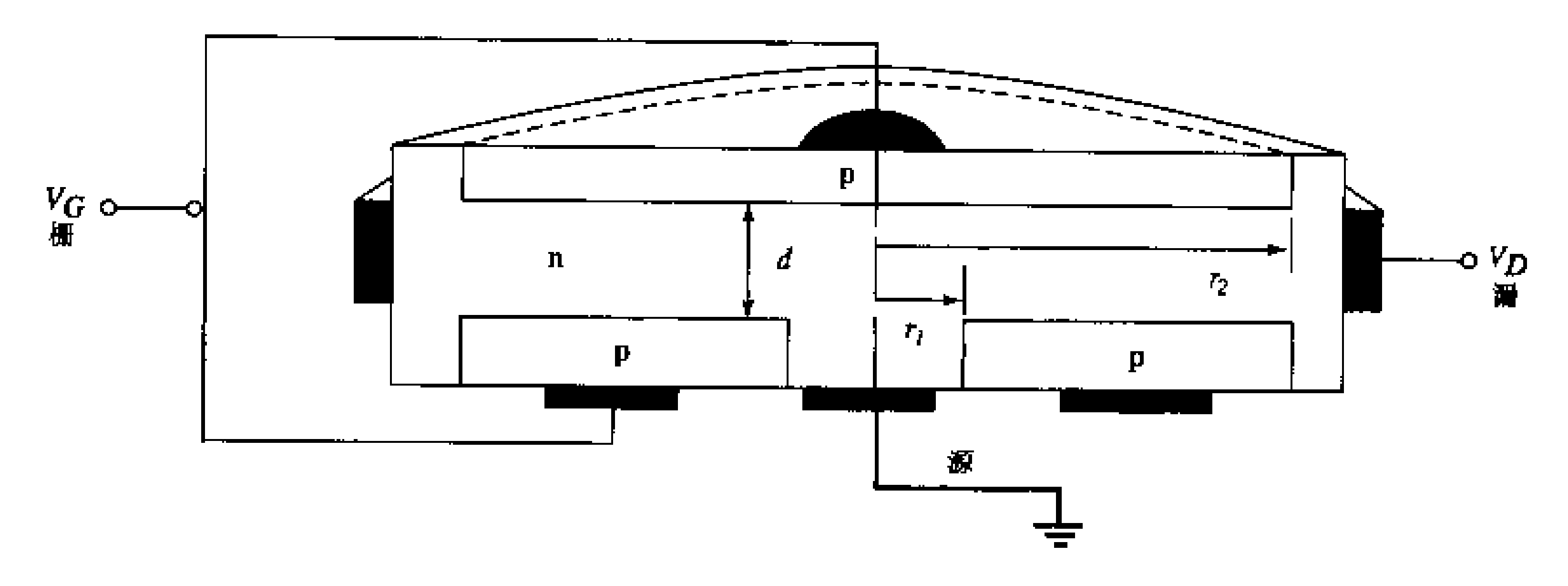


图 P4.16

第5章 pn结中的电流

迄今为止,我们已经讨论了反偏pn结。只要外加电压没有超过结的击穿电压,pn结上只有小电流流过。我们还没有考虑低偏压下的电流情况,只是集中讨论了pn结中阻止多数载流子迁移的势垒,以及耗尽区宽度随外加电压的变化。

本章主要讨论正偏和反偏条件下流过pn结的电流。这部分内容很重要,不仅可以掌握结型二极管特性,而且可以了解结型晶体管的基本概念。分析电流特性的第一步需要导出自由载流子的连续性方程,也就是在半导体中无限小的空间内考虑影响载流子浓度的各种机制的方程。为了在方程中描述重要的产生与复合过程,需要比第1章更细致地考虑一些基本的物理过程。推导出包含了产生与复合项的连续性方程后,就可以表征偏置条件下pn结准中性区少数载流子的分布情况。我们将详细讨论两种特殊的pn结,得到理想二极管特性。然后,我们将这一结果用于实际的硅二极管,并讨论空间电荷区的产生与复合。稳态电流电压关系的物理模型将有助于理解电荷存储效应和二极管瞬变特性。在最后章节我们将讨论pn结在集成电路中的作用,给出理论的实际用途。

5.1 连续性方程

为了讨论pn结电流,我们先写出流入和流出空间无限小体积的自由载流子通量的方程。对半导体中的多数载流子和少数载流子都可以写出这种连续性方程,其中求解少数载流子的连续性方程在很多实际器件中有特殊的重要性。

首先推导电子的一维连续性方程。考虑位于x处厚为dx的无限小薄层,如图5.1所示,由于净的流入该体积的电子以及薄层中净的载流子的产生,薄层中电子的数量会增加。电子增加的总速率等于以下四项的代数和:

(1)进入薄层的电子数,负;

(2)流出薄层的电子数,正;

(3)电子的产生率,负;

(4)电子的复合率。

薄层每一侧的电流除以电子电荷可得到前两项分量,后两项分量分别用G和R表示,那么薄层中电子浓度随时间的变化率为

$$\frac{\partial n}{\partial t}A\,dx = \left(\frac{J_n(x)}{-q} - \frac{J_n(x + dx)}{-q}\right)A + (G_n - R_n)A\,dx \tag{5.1.1}$$

式中A是薄层截面积,G_n和R_n分别代表单位体积内电子的产生率和复合率。将等式右端第二项用台劳级数展开

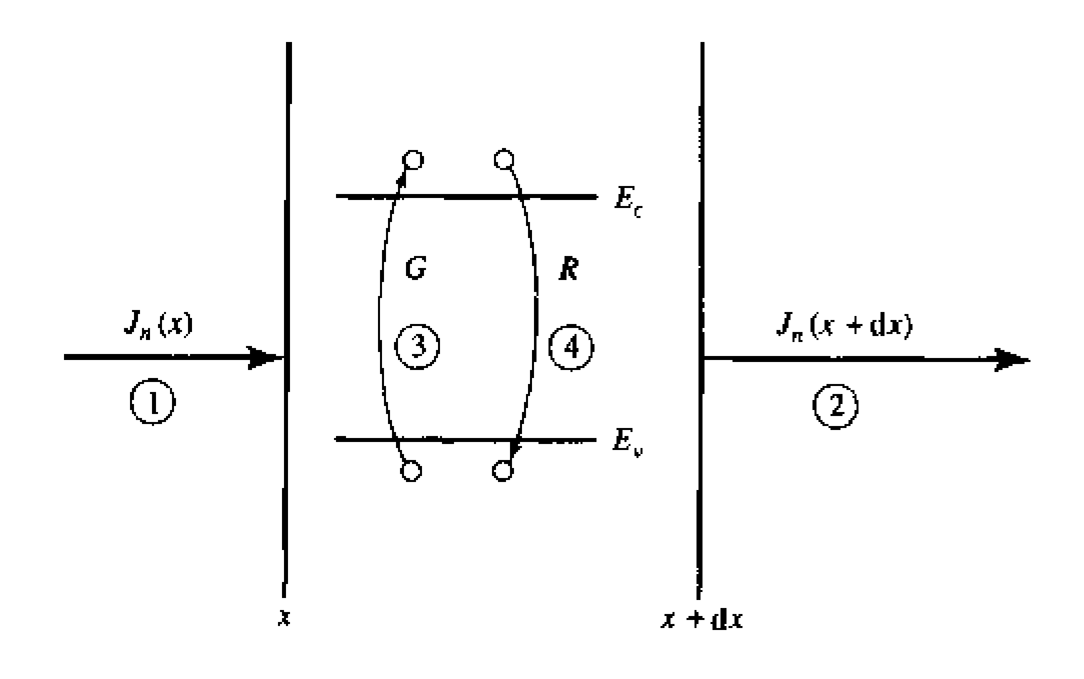


图5.1 在厚为 $\mathrm{d}x$ 的无限小薄层中电子浓度的增加与净的流入薄层的电子流以及产生率与复合率之差有关

$$J_n(x+\mathrm{d}x)=J_n(x)+\frac{\partial J_n}{\partial x}\mathrm{d}x+\cdots \tag{5.1.2}$$

得到基本的电子连续性方程

$$\frac{\partial n}{\partial t}=\frac{1}{q}\frac{\partial J_n}{\partial x}+(G_n-R_n) \tag{5.1.3a}$$

空穴的连续性方程与等式(5.1.3a)类似,不同的是右侧的第一项改变了符号,因为电荷是空穴

$$\frac{\partial p}{\partial t}=-\frac{1}{q}\frac{\partial J_p}{\partial x}+(G_p-R_p) \tag{5.1.3b}$$

为了求解等式(5.1.3),必须将右侧的表达式与载流子浓度 n 和 p 联系起来。对于电流项很简单,因为等式(1.2.21)和(1.2.22)中 J_n 和 J_p 是用载流子浓度表示的。将这些等式代入得到

$$\frac{\partial n}{\partial t}=\mu_n n(x)\frac{\partial \mathscr{E}(x)}{\partial x}+\mu_n\mathscr{E}(x)\frac{\partial n(x)}{\partial x}+D_n\frac{\partial^2 n(x)}{\partial x^2}+(G_n-R_n) \tag{5.1.4a}$$

和

$$\frac{\partial p}{\partial t}=-\mu_p p(x)\frac{\partial \mathscr{E}(x)}{\partial x}-\mu_p\mathscr{E}(x)\frac{\partial p(x)}{\partial x}+D_p\frac{\partial^2 p(x)}{\partial x^2}+(G_p-R_p) \tag{5.1.4b}$$

式中假定迁移率 μ 和扩散系数 D 与 x 无关。虽然这个假设在许多情况下不成立,但是等式(5.1.4)还是包含了主要的物理效应,并且可通过数值方法进行更精确的计算。

如果所考虑区域的电场等于零,或者可忽略,等式(5.1.4)中右侧的前两项就可以忽略,分析将大大简化。即使电场不能忽略,等式(5.1.4)中的某些项也不重要。例如,如果电场是常数,等式(5.1.4)中的第一项就等于零。如4.1节所看到的,指数掺杂分布的半导体中电场就为常数。一般不必处理等式(5.1.4)中的所有项。

在求解连续性方程之前,有必要回顾微积分中的一些基本概念。连续性方程(等式(5.1.4))是一个偏微分方程,是时间和位置的函数,所以有无限多个解。给定的问题必须有边界条件和初始条件才能得到确定的解。如果只需求稳态解,连续性方程将简化成普通的微分方程。在这种情况下,等式左侧与时间相关的项等于零,只留下与位置有关的项。

5.2 产生与复合

为了得到用自由载流子浓度表示的电子的产生率与复合率的表达式，需要在第1章的基础上，进一步讨论半导体物理概念。第1章指出，价带中的电子可被热激发到导带中，在价带中留下一个空穴，空穴和电子都对电导有贡献。其反过程为电子直接跃迁到价带中。热平衡时，电子产生率等于复合率。这是自由载流子产生与复合的一种途径；当然，也有其他的产生-复合途径。

所有半导体中都会发生电子在价带和导带之间的直接跃迁。在许多化合物半导体中，直接跃迁是最重要的产生-复合机制，例如，由元素周期表的Ⅲ族和Ⅴ族元素构成砷化镓(GaAs)和磷化铟(InP)。然而，由于能带结构的关系，硅和锗中一般不发生直接跃迁，除非空穴和电子的浓度非常高。硅和锗的导带底的电子"动量"不为零，而价带顶的空穴动量为零，所以，为保证能量和动量守恒，如果没有晶格(声子)的参与，不可能发生直接跃迁。因此，硅和锗中跨越禁带的直接跃迁需要三个粒子同时参与，即电子、空穴和代表晶格作用的声子(参考1.2节)。

三个粒子发生相互作用的几率比两个粒子的小，例如，一个自由载流子和一个声子之间的两粒子相互作用，电子或空穴可以在声子的作用下跃迁到局域的允许态。事实上，能量位于E_v和E_c之间的局域态总是存在的，因为晶体中存在晶格缺陷，或者更常见的杂质原子。而且，由于硅和锗中的局域态浓度很高，所以在产生与复合过程中起主导作用。这些局域态起到复合中心[1]的作用。例如，在复合过程中，复合中心俘获一个导带电子，然后这个电子跃迁到价带中的空状态，与一个空穴复合。

5.2.1 局域态：俘获和发射

自由载流子与局域态相互作用的四种过程如图5.2所示。图中局域态位于禁带中能量E_t处，浓度为N_t。图中所示的是受主型局域态，即能级空时呈中性，被占据时带负电，图中描述的过程同样适用于施主型局域态。

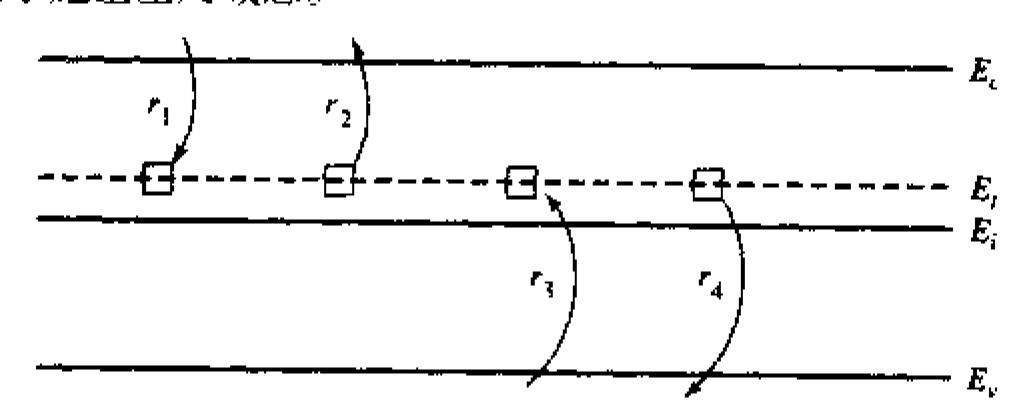


图5.2 自由载流子与局域态作用的四个过程：r_1为俘获电子过程；r_2为发射电子过程；r_3为俘获空穴过程；r_4为发射空穴过程。图中所示局域态为受主型，能量E_t位于禁带中

俘获电子过程是指一个电子从导带坠落到空的局域态中。这一过程发生的几率与导带电

[1] 因为这些态对自由载流子的产生与复合过程都是中间过渡状态，所以被称为产生-复合中心，简称为"复合中心"。

子浓度 n、空的局域态密度以及电子被俘获的几率成正比。空的局域态密度等于总局域态密度 N_t 乘以 $[1-f(E_t)]$，其中 $f(E_t)$ 为局域态被占据的几率。热平衡时，f 等于 f_D，f_D 就是等式(1.1.18)给出的 Fermi 分布函数。非平衡时，f 不等于 f_D，这里不再进一步讨论。

单位时间一个电子被局域态俘获的几率等于电子热运动速度 v_{th} 与电子俘获截面 σ_n 的乘积。电子俘获截面表示局域态俘获电子的能力。$v_{th}\sigma_n$ 可以被形象地看做是单位时间，具有横截面为 σ_n 的粒子扫过的体积。如果局域态位于这一体积中，电子就被俘获。对给定类型的局域态，俘获截面通常由实验确定。杂质金和铁的有效电子俘获截面的典型值大约为 10^{-15} cm^2[1]。铍的俘获截面非常大，约为 $10^{-10}cm^2$[2]。通过以上讨论，可得出局域态俘获电子的总速率(电子俘获率)为

$$r_1 = n\{N_t[1 - f(E_t)]\}v_{th}\sigma_n \tag{5.2.1}$$

发射电子过程与俘获电子过程相反，指的是电子从局域态激发到导带。这一过程发生的速率等于被电子占据的局域态密度 $N_t f(E_t)$ 乘以电子的跃迁几率 e_n

$$r_2 = [N_t f(E_t)]e_n \tag{5.2.2}$$

热平衡下的电子产生率等于等式(5.2.1)表示的电子俘获率，即俘获和发射载流子的速率必须相等，$f(E)$ 由 Fermi 分布函数 $f_D(E)$ 给出(等式(1.1.18))。这样，可以写出

$$r_1 = r_2 = nN_t[1 - f_D(E_t)]v_{th}\sigma_n = N_t f_D(E_t)e_n \tag{5.2.3}$$

和

$$e_n = v_{th}\sigma_n n_i \exp\left(\frac{E_t - E_i}{kT}\right) \tag{5.2.4}$$

一般用等式(5.2.4)描述等式(5.2.2)中的 e_n。从等式(5.2.2)可看出，当局域态能量接近于导带边时，电子从局域态发射到导带的可能性增加，因为此时的 $E_t - E_i$ 增大。

局域态和价带之间的相互作用也有相应的关系式。空穴俘获率与被电子占据的局域态密度 $N_t f(E_t)$、空穴密度以及跃迁几率成正比。这一几率可表示为空穴热运动速度 v_{th} 和空穴俘获截面 σ_p 的乘积。这样

$$r_3 = [N_t f(E_t)]pv_{th}\sigma_p \tag{5.2.5}$$

发射空穴过程指的是一个电子从价带激发到空的局域态。与前面的发射电子过程类似，空穴产生率可由下式给出

$$r_4 = \{N_t[1 - f(E_t)]\}e_p \tag{5.2.6}$$

热平衡下空穴的跃迁几率 e_p 可在用 σ_p 表示出，其中 $r_3 = r_4$

$$e_p = v_{th}\sigma_p n_i \exp\left(\frac{E_i - E_t}{kT}\right) \tag{5.2.7}$$

与等式(5.2.4)类似，当能量接近价带边时，空穴从局域态向价带的发射几率增大。

在利用等式(5.2.1)、(5.2.2)、(5.2.5)和(5.2.6)研究导带和价带通过局域态 E_t 相互作用的动力学问题之前，有必要定性考虑它们的物理含义。首先，在热平衡时，$r_1 = r_2$，$r_3 = r_4$，因为热平衡要求每个过程和它的反过程相互抵消。非平衡时，$r_1 \neq r_2$ 且 $r_3 \neq r_4$。为了得到这些速率，假定如下情况：n 型半导体中的空穴数量突然增加并超过其热平衡值，这将导致 r_3 增加，结果导致 r_4 和

r_1同时增加（这两个过程都能消除E_t处的空穴）。如果大部分的空穴通过E_t与导带中的电子复合，即r_1过程，局域态是有效的复合中心。如果大部分的空穴通过r_4过程从能级E_t转移到价带，那么E_t位置就是有效的空穴陷阱。给定的局域态通常只通过一种方式起作用：或者作为陷阱，或者作为复合中心。如果局域态位于带隙中心，则可能成为有效的复合中心；如果接近于能带边，那么更可能是载流子陷阱，而不是复合中心。下面将更加详尽地论述复合中心的作用。

5.2.2 Shockley-Hall-Read 复合[+]

通过局域态或复合中心来描述产生与复合过程的方程，最初是由 Shockley、Read[3] 和 Hall[4] 推导出来的，所以这一过程通常被称做 Shockley-Hall-Read(或者 SHR)复合。根据 SHR 模型，当半导体处于非平衡态时，位于复合中心的总的电子和空穴数量与平衡态相比，并不发生大的改变。原因是复合中心虽然能快速俘获多数载流子(附近有很多)，但必须等待少数载流子到来才能复合，所以，复合中心上几乎总是充满了多数载流子，无论是热平衡还是非平衡状态。

为了解释这一现象，考虑一个典型的情况：分析 n 型半导体中受主型复合中心。热平衡时，Fermi 能级更接近E_c，比复合中心的能量要高，因此，复合中心实际上填满了电子，r_1和r_2比r_3或r_4大很多。当平衡态被较小的激励破坏时，空穴和电子增加的数量相同，但总的电子浓度改变不大，而空穴浓度却发生了较大的变化。在这种情况下，r_1将略微超过r_2以平衡r_3所表示的空穴俘获率的增加。这样，被占据的局域态数量几乎保持常数，净的电子俘获率r_1-r_2等于净的空穴俘获率，也就是净复合率，用符号U表示

$$U \equiv R_{sp} - G_{sp} = r_1 - r_2 = r_3 - r_4 \tag{5.2.8}$$

式中的下角标 sp 代表自发(Spontaneous)，就是说复合和产生仅仅是对偏离热平衡的一种响应[2]。将r_1到r_4的表达式代入等式(5.2.8)，消去f，解出U

$$U = \frac{N_t v_{th}\sigma_n\sigma_p(pn - n_i^2)}{\sigma_p\left[p + n_i\exp\left(\dfrac{E_i - E_t}{kT}\right)\right] + \sigma_n\left[n + n_i\exp\left(\dfrac{E_t - E_i}{kT}\right)\right]} \tag{5.2.9a}$$

$$= \frac{(pn - n_i^2)}{\tau_{no}\left[p + n_i\exp\left(\dfrac{E_i - E_t}{kT}\right)\right] + \tau_{po}\left[n + n_i\exp\left(\dfrac{E_t - E_i}{kT}\right)\right]} \tag{5.2.9b}$$

式中$\tau_{no} = (N_t v_{th}\sigma_n)^{-1}$，$\tau_{po} = (N_t v_{th}\sigma_p)^{-1}$。

等式(5.2.9)表明，如果pn乘积超过n_i^2，则U为正，存在净复合。反之，如果pn乘积小于n_i^2，则U为负，存在净产生。$(pn - n_i^2)$项相当于自由载流子数目非平衡条件下的恢复“力”。

如果电子和空穴俘获截面相同，等式(5.2.9)中U与复合中心能级的关系更简单。当$\sigma_p = \sigma_n \equiv \sigma_0$时，定义$\tau_0 \equiv (N_t v_{th}\sigma_0)^{-1}$，那么

$$U = \frac{(pn - n_i^2)}{\left[p + n + 2n_i\cosh\left(\dfrac{E_t - E_i}{kT}\right)\right]\tau_0} \tag{5.2.10}$$

② 与自发复合和产生相反的是那些由激发所导致的跃迁，例如，辐射。

与复合中心能级的关系包含在等式(5.2.10)中的双曲余弦项中。这一项关于 $E_t = E_i$ 对称,反映了复合中心俘获空穴和电子的对称性。分母在 $E_t = E_i$ 时具有最小值,表明复合中心靠近带隙中心时,U 具有最大值。图 5.3 中的实线代表 n 型半导体的复合情形,曲线根据等式(5.2.10)绘出,并按 U 的最大值归一化,曲线横坐标是 $(E_t - E_i)/kT$。与本图相关的其他条件是:$p < n$,$n = 10^{16}\text{cm}^{-3}$,$(pn - n_i^2) = 1.5 \times 10^{31}\text{cm}^{-6}$,$\tau_0 = 10^{-7}\text{s}$。我们将发现这些值的选取对正向偏置的 pn 结准中性区是合理的。图 5.3 中的虚线代表产生率 U(按最大值归一化)与 $(E_t - E_i)/kT$ 的关系,其他已知条件如下:$(pn - n_i^2) = -2.1 \times 10^{20}\text{cm}^{-6}$,$p$ 和 n 远小于 n_i,$\tau_0 = 10^{-7}\text{s}$。反偏 pn 结耗尽区的中心附近可能满足这些条件。图 5.3 表明半导体耗尽区的产生率与复合中心能量的关系比未耗尽区复合率的依赖关系更明显。这是因为,当载流子浓度都很小时,为了使载流子在复合中心与导带和价带之间的转移达到平衡,复合中心起到的作用更大。然而,无论是产生还是复合,最有效的复合中心是接近于 E_i 的 E_t。例如,金和铜将引入两种有效的复合中心。金和铜在硅中的 $(E_t - E_i)$ 的值分别是 0.03eV 和 0.01eV[15]。表 1.4 给出了许多元素在硅中的能级。

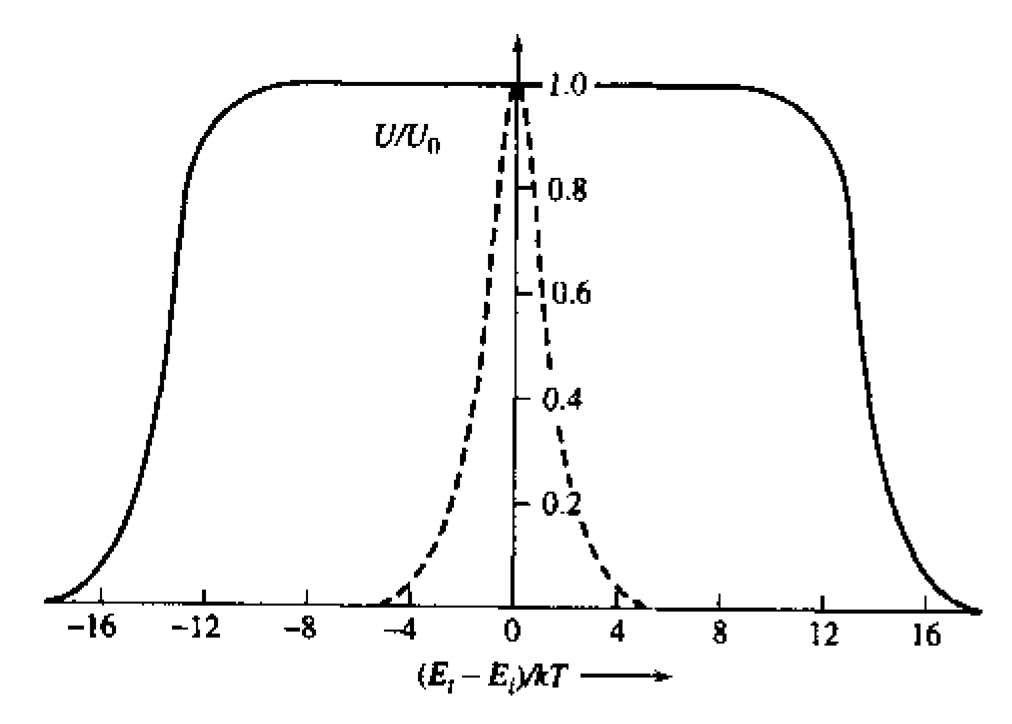


图5.3 复合率(实线)和产生率(虚线)与 $(E_t - E_i)/kT$ 的关系。利用等式(5.2.10)及文中给出的数值绘出,并用 $E_t = E_i$ 时的 U 值归一化

等式(5.2.9)和(5.2.10)是分析 SHR 复合得到的主要结论。它们表明,通过复合中心的净复合率 U 是自由载流子浓度的函数,同时也与复合中心的性质有关。对于特定的问题,方程经常可以被简化。根据这些方程对材料和器件性能的预测通常需要实验验证。

5.2.3 过剩载流子寿命

为了理解净复合率 U 的物理意义,考虑没有电流流动的半导体,热平衡被突然产生的等量的过剩电子和空穴破坏。通过过剩载流子自发衰减,半导体又回到热平衡状态。解连续性方程(等式(5.1.3))可得到这种情况的过剩电子浓度与时间的函数关系。我们用一个实际情况下通常可以满足的假设:小注入。在这个条件下,外在的扰动不会显著改变平衡时总的自由载流子浓度。如果用 n' 表示过剩电子浓度,用 p' 表示过剩空穴浓度,那么小注入表明 n' 和 p' 都远小于 $(n_o + p_o)$,这里 n_o 和 p_o 分别表示热平衡时半导体中的电子和空穴浓度。利用这些定

义得到，$n' \equiv n - n_o$，$p' \equiv p - p_o$，并且 $n' = p'$[3]。

如果 $\sigma_p = \sigma_n$，那么复合率 U 由等式(5.2.10)给出，连续性方程(5.1.3a)可以写成

$$\frac{dn'}{dt} = G - R = -U = \frac{-(n_o + p_o)n'}{\left(n_o + p_o + 2n_i \cosh\left[\frac{E_t - E_i}{kT}\right]\right)\tau_0} \tag{5.2.11}$$

我们发现解出的过剩载流子浓度 n' 随时间指数衰减

$$n'(t) = n'(0)\exp(-t/\tau_n) \tag{5.2.12}$$

其中寿命 τ_n 由下式给出

$$\tau_n = \left[\frac{n_o + p_o + 2n_i \cosh\left(\frac{E_t - E_i}{kT}\right)}{(n_o + p_o)}\right]\tau_0 \tag{5.2.13}$$

正如前一节指出的，$(E_t - E_i)$ 相对小的复合中心比较有效，所以等式(5.2.13)分子上第三项与前两项之和相比较可以忽略，于是等式(5.2.13)简化为

$$\tau_n = \tau_0 = \frac{1}{N_t v_{th} \sigma_o} \tag{5.2.14}$$

和

$$U = \frac{n'}{\tau_n} \tag{5.2.15}$$

等式(5.2.14)表明小注入下过剩载流子的寿命与通过复合中心复合的多数载流子浓度无关。我们可以利用复合的动力学过程从物理上理解这一现象。例如，p 型半导体中，因为 $E_f < E_t$（假定陷阱靠近带隙中部），大部分的复合中心是空的，没有被电子占据，所以复合过程受到俘获导带电子的限制。一旦电子被复合中心俘获，由于价带中空穴很多，复合中心能快速俘获价带中的一个空穴。这样，复合过程主要受到复合中心对少数载流子的俘获速率的限制，而对多数载流子的数量不敏感。

少数载流子寿命的变化范围可以很宽，主要取决于半导体中复合中心的类型和浓度。探测器要求少数载流子的寿命长，可以通过特殊的处理在硅中得到毫秒量级甚至更长的少子寿命。集成电路中的少子寿命的典型值分布在不到一个微秒至数百微秒的范围内。

Auger 复合[+]　前一节讨论了 Shockley-Hall-Read 复合，即过剩载流子通过复合中心的复合。其过程是一个过剩载流子被复合中心俘获，直到一个具有相反导电类型的载流子到达复合中心时发生复合。SHR 复合在载流子浓度较低或者中等时起主要作用，因为载流子更倾向于与复合中心发生作用而不是低浓度的可动载流子。

然而，载流子浓度较高时，电子和空穴的直接相互作用将导致 Auger 复合[4]。Auger 复合中，导带中的一个电子落入价带中的空状态（空穴），这一过程释放出的能量被另一个载流子吸收，以保证动量守恒。Auger 复合是 4.4 节所讨论的雪崩产生电子-空穴对的反过程，发生雪

3　因为电子和空穴以相同的速率增加或者减少，过剩载流子浓度 n' 和 p' 相等。

4　以物理学家 P. Auger 命名，读音为"oh-zhay"。

崩时入射载流子的能量和动量激发出电子-空穴对。

Auger 复合过程有三个载流子参与:复合的电子和空穴,以及接受能量的载流子。发生 Auger 复合要求三个载流子相互作用的概率必须足够大,所以,只会发生在高掺杂材料中,或者过剩载流子的浓度非常大时。在 n 型材料中,是两个电子和一个空穴发生作用,而在 p 型材料中,是两个空穴和一个电子相互作用。因为有两个多数载流子参与 Auger 复合,复合率 U 将与多数载流子浓度的平方成正比。Auger 复合的复合率 U_A 由下式给出

$$U_A = R_A - G_A = c_n n(pn - n_i^2) + c_p p(pn - n_i^2) \tag{5.2.16}$$

式中 c_n 和 c_p 是 Auger 复合系数。

对于 p 型重掺杂材料中的电子复合,Auger 复合寿命的倒数可以写成

$$\frac{1}{\tau_A} = c_n N_a^2 \tag{5.2.17}$$

硅中电子的复合系数 c_n 大约是 $1 \times 10^{-31} \mathrm{cm^{-6} s^{-1}}$,n 型材料中空穴的 Auger 复合寿命大约只有 p 型材料中电子寿命的 1/3 到 1/2。既考虑 SHR 复合,又考虑 Auger 复合,那么有效寿命 τ 为

$$\frac{1}{\tau} = \frac{1}{\tau_{\mathrm{SHR}}} + \frac{1}{\tau_A} \tag{5.2.18}$$

因为 n 型区掺杂通常比 p 型区高(例如,双极晶体管的发射区,或者 n 沟道 MOSFET 的源区和漏区),所以 Auger 复合通常在 n 型材料中更为重要。

表面复合† 迄今为止,我们所考虑的产生-复合中心是在整个半导体材料上均匀分布的。然而,3.5 节指出,半导体表面存在大量的局域态,其能量位于禁带中。通常由平面工艺制备的器件的半导体表面存在二氧化硅钝化层,与可能成为表面态的悬挂键结合,阻止外界原子的进入。钝化的氧化层可以将表面态密度从 $10^{15} \mathrm{cm^{-2}}$ 减少到 $10^{11} \mathrm{cm^{-2}}$。即使存在表面钝化层,表面态除了提供体内的局域态外,依然能提供产生-复合中心。因为许多实际的半导体器件的性质受到表面产生-复合的影响,所以有必要简单地讨论这个问题。

表面产生-复合的动力学和那些体内的复合中心是类似的,只有一点不同:体内复合中心考虑的是体密度 $N_t(\mathrm{cm^{-3}})$,而表面必须考虑表面复合中心的面密度 $N_{st}(\mathrm{cm^{-2}})$[5]。尽管表面复合中心 N_{st} 可以分布在厚约几个原子层的范围内,但由于半导体表面微观结构不好定义,所以用数量相等的表面态来等效。与等式 5.2.9 类似,写出单位面积表面复合率 U 的表达式

$$U_s = \frac{N_{st} v_{th} \sigma_n \sigma_p (p_s n_s - n_i^2)}{\sigma_p\left[p_s + n_i \exp\left(\frac{E_i - E_{st}}{kT}\right)\right] + \sigma_n\left[n_s + n_i \exp\left(\frac{E_{st} - E_i}{kT}\right)\right]} \tag{5.2.19}$$

其中下角标 s 表示表面附近,E_{st} 表示表面产生-复合中心的能量。为了强调表面复合的重要性,我们考虑带隙中心附近的最有效的复合中心。假定电子和空穴俘获截面相等,等式(5.2.19)可简化为

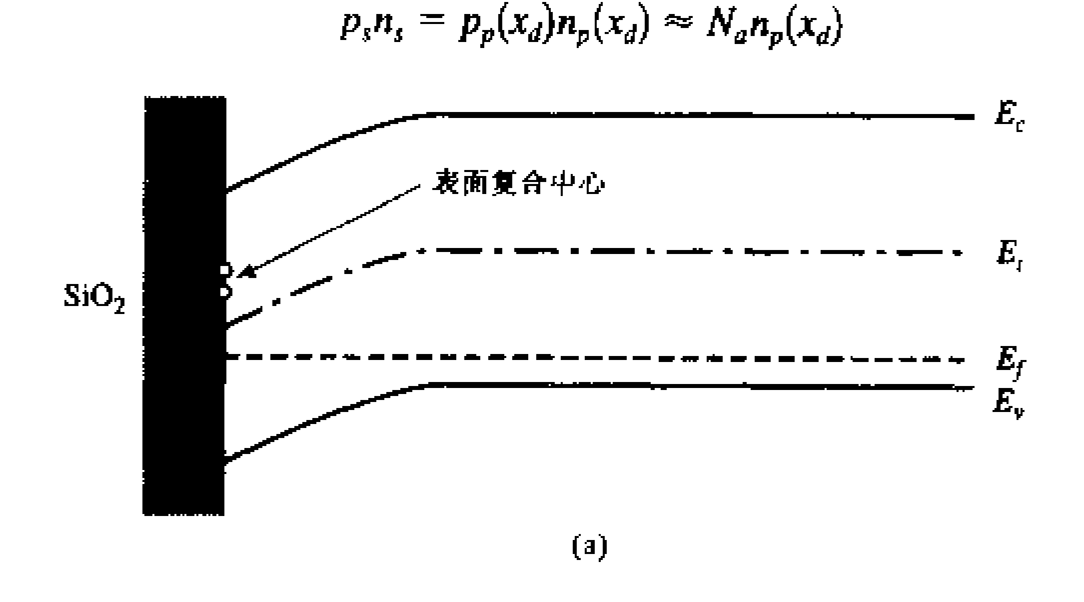
$$U_s = N_{st} v_{th} \sigma \frac{(p_s n_s - n_i^2)}{p_s + n_s + 2n_i \cosh\left(\frac{E_{st} - E_i}{kT}\right)} \tag{5.2.20}$$

---

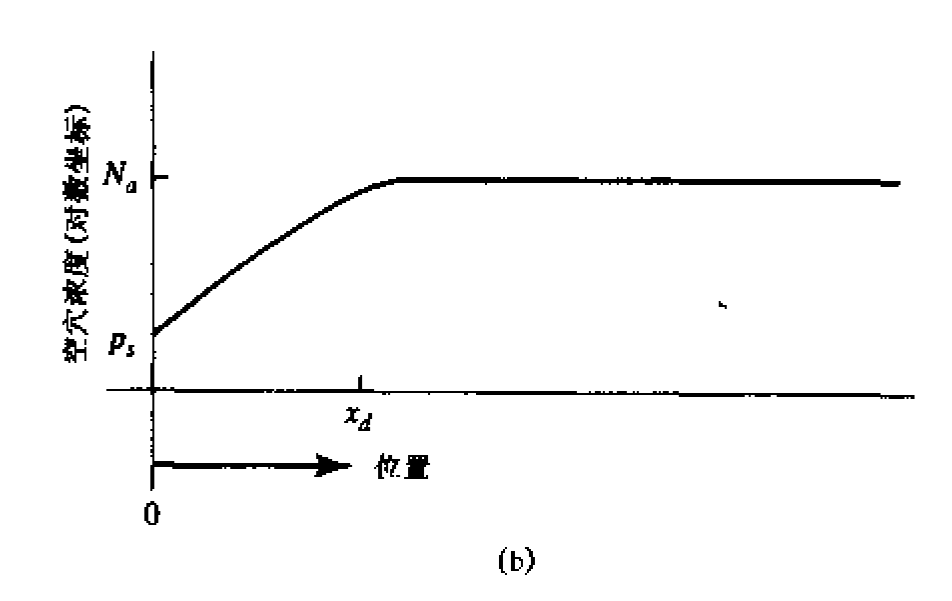
[5] N_{st} 表示作为产生-复合中心的界面陷阱态密度。我们将在 8.5 节中进一步讨论。

第 3 章指出半导体表面和体内电势通常不同,这样表面载流子浓度与体内的中性区不同。即使表面有氧化钝化层存在,通常在半导体表面附近也会形成空间电荷区,如图 5.4 所示的 p 型硅。第 2 章指出,二氧化硅界面上的杂质分凝,将导致硅表面与体内相比,是更弱的 p 型或者更强的 n 型。如果假设 pn 乘积在整个空间电荷区保持为常数,那么 p 型半导体表面上的乘积 p_sn_s 可以根据空间电荷区中性区边界的数值表示[⑥]

$$p_sn_s = p_p(x_d)n_p(x_d) \approx N_a n_p(x_d) \tag{5.2.21}$$



(a)



(b)

图 5.4 (a)覆盖有钝化氧化层的 p 型硅表面能带图;(b)表面空穴浓度分布

等式(5.2.20)可以写成

$$U_s = N_{st}v_{th}\sigma\frac{N_a[n_p(x_d) - n_{po}]}{(p_s + n_s + 2n_i)} = N_{st}v_{th}\sigma\frac{N_a}{(p_s + n_s + 2n_i)}n_p'(x_d) \tag{5.2.22}$$

式中假定 $E_{st} \approx E_i$。等式(5.2.22)中表面复合率 U_s 用表面空间电荷区过剩少数载流子(电子)浓度 n_p' 来表示。

等式(5.2.22)中右侧 n_p' 的系数通常被定义成参数 s,描述表面复合过程的特性

$$s = N_{st}v_{th}\sigma\frac{N_a}{(p_s + n_s + 2n_i)} \tag{5.2.23}$$

s 值取决于表面产生-复合中心的物理特性、浓度以及表面电势。如果表面区可动载流子耗尽,

---

⑥ 下标 p 和 n 分别指 p 型和 n 型材料的载流子浓度。

那么p_s和n_s较小，s较大。如果表面是中性的，$p_s \sim N_a$，s较小，且

$$s = s_o = N_{st} v_{th} \sigma \tag{5.2.24}$$

式中下标o表示表面和体内电势相同，也就是表面区是中性的。s对表面势的依赖关系在硅集成电路中非常重要。

s的量纲是cm s^{-1}，因此s通常被称为表面复合速度，尽管它并不直接与速度相关。比较等式(5.2.24)和(5.2.14)中的少数载流子寿命，可以这样从物理上理解s：即s和表面过剩载流子复合率的关系与$1/\tau$和体内复合率的关系相同。

例题　表面复合速度

已知室温下的n型硅电阻率$\rho_n = 0.025\Omega \cdot$ cm，整个SiO_2钝化层处于光照下，所以空穴和电子浓度比热平衡值高。假设热平衡时表面电子浓度$n_s = 10^{16}$ cm^{-3}，载流子的光致产生速度是10^{14} cm^{-2}s^{-1}，在光照下，表面空穴浓度增加到$p_s = 10^{10}$ cm^{-3}。

(a)当50%的过剩载流子在表面复合时，求表面复合速率s。

(b)如果表面复合是由浓度为$N_{st} = 10^{11}$ cm^{-2}的表面复合中心引起的，求由(a)求出的复合速度引起的俘获截面σ。

解：　首先根据图1.15或者表4.1，由$\rho_n = 0.025\Omega \cdot$ cm查出$N_d = 10^{18}$ cm^{-3}。因为$n_s = 10^{16}$ cm^{-3}，即使处于光照下，半导体表面仍是多数载流子轻耗尽。利用等式(1.1.13)，得到表面空穴平衡浓度

$$p_s = n_i^2 / n_s = 2.1 \times 10^4 \text{ cm}^{-3}$$

所以，表面过剩空穴浓度为

$$p_s' = 10^{10} - 2.1 \times 10^4 \approx 10^{10} \text{ cm}^{-3}$$

表面复合速度正好等于光致产生速度的一半，这样

$$U_s = s p_s' = s \times 10^{10} = 0.5 \times 10^{14} \text{ cm}^{-2}\text{s}^{-1}$$

所以，$s = 5000$ cm s^{-1}，这就是(a)的答案。对于(b)，利用适用于n型硅的等式(5.2.23)，以及等式(5.2.24)，计算得到

$$s_o = s\frac{(p_s + n_s + 2n_i)}{N_d} \approx 5000 \frac{10^{16}}{10^{18}} = 50 \text{ cm s}^{-1}$$

取$N_{st} = 10^{11}$ cm^{-2}，$v_{th} = 10^7$ cm s^{-1}(1.2节)，得到$\sigma = 5 \times 10^{-17}$ cm^2。

5.3　pn结电流电压特性

根据连续性方程(等式(5.1.3))和从SHR产生-复合模型推导出来的过剩载流子概念，我们可以求出pn结在外加偏置下的电流表达式。准中性区连续性方程的解表明载流子浓度是与位置和时间相关的。根据载流子浓度可求出的载流子电流，其表达式可直接从等式(1.2.21)和(1.2.22)得到。总电流通常由四个分量之和组成：空穴和电子漂移电流、空穴和电子扩散电流。

考虑连接在电压源上的pn结二极管，n型区接地，p型区相对于地的电压是V_a。二极管

具有恒定的横截面A，横截面图如图5.5所示。在没有光照的情况下，二极管中的载流子浓度仅受外加电压的影响。外加电压V_a部分降落在准中性区，部分降落在结上。因为降落在准中性区的电压满足欧姆定律（电流乘以电阻），在低的或中等电流下的IC器件中该电压是很小的。但是，对于横截面较小的器件，如在ULSI的某些电路中，这个欧姆压降会限制器件的性能。

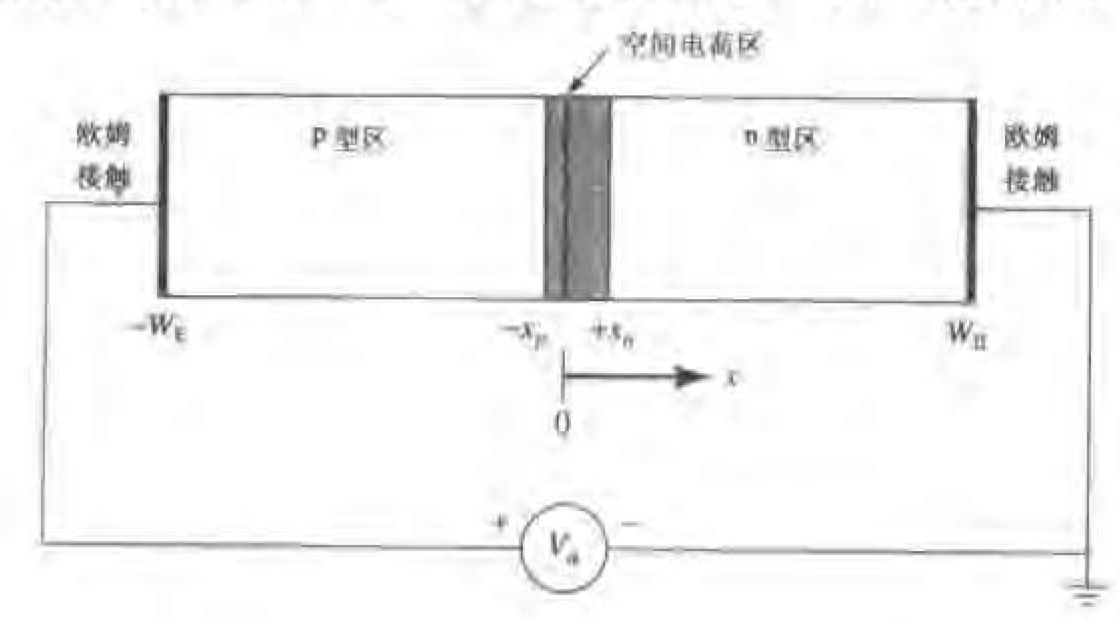


图5.5 用于讨论电流与电压关系的pn结二极管结构。图中指出了坐标轴和偏压方向。假定横截面A是均匀的

我们的分析忽略欧姆压降，假设V_a全部落在结上。于是，外加偏压下，总的结电压就是$\phi_i - V_a$，其中ϕ_i是内建电压。如果V_a为正（正偏），外加电压将降低结上多数载流子扩散的势垒。反过来，势垒的减小使得空穴从p型一侧向n型一侧有净的迁移，同时电子从n型一侧向p型一侧也有净的迁移。这些载流子迁移到准中性区后，变成少数载流子，被从欧姆接触电极进入到准中性区的多数载流子快速中和了。这种注入载流子的中和就是我们在第1章中曾经讨论过的介电弛豫。一旦少数载流子通过空间电荷区注入，它们趋向于从结向中性区扩散。

如果V_a是负的（反偏），多数载流子扩散的势垒高度增加，平衡被破坏，靠近结空间电荷区的少数载流子趋向于耗尽。多数载流子浓度的减少过程类似于介电弛豫。

这些简单的说明表明应该注意少数载流子浓度的特殊作用，因为它们决定了pn结中电流的大小。多数载流子仅仅扮演了提供注入的少数载流子电流，或者在准中性区中和电荷的作用，可以认为多数载流子受少数载流子的支配。因此，我们将在准中性区求解少数载流子连续性方程。

5.3.1 边界少数载流子浓度

为了将这些解写成一个有用的形式，必须将少数载流子浓度的边界值与外加电压V_a联系起来。最直接的方式是做另外的两个假设：第一，外加偏压导致小注入；第二，外加电压足够小，以至于结区中多数载流子和少数载流子之间的平衡没有被显著破坏。5.2节已经讨论过小注入。简单地说，它表明在外加偏压下，准中性区边界上的多数载流子浓度不会显著改变。第二条假设允许在结上使用等式(4.1.9)，电势差是$\phi_i - V_a$。

当V_a较小（$|V_a| \ll \phi_i$）时这两个假设都成立。在高偏压下是否有效就需要慎重考虑了，我们将在后面继续分析。

根据小注入假设，无论是在平衡或者偏置下，靠近pn结的n型准中性区边界上的电子浓

度等于掺杂浓度,和第4章中一样,这个边界位置为x_n(图5.5),用下标o表示热平衡态。

同样,准中性p型区的边界在$-x_p$处,该处的空穴浓度等于平衡时的受主掺杂浓度,有外加偏置时也成立。用等式表示这些结论,得到

$$n_{po}(-x_p) = n_{no}(x_n)\exp\left(\frac{-q\phi_i}{kT}\right)$$
$$= N_d(x_n)\exp\left(\frac{-q\phi_i}{kT}\right) \tag{5.3.1}$$

$$p_{no}(x_n) = p_{po}(-x_p)\exp\left(\frac{-q\phi_i}{kT}\right)$$
$$= N_a(-x_p)\exp\left(\frac{-q\phi_i}{kT}\right) \tag{5.3.2}$$

$$n_p(-x_p) = N_d(x_n)\exp\left[\frac{-q(\phi_i - V_a)}{kT}\right] \tag{5.3.3}$$

$$p_n(x_n) = N_a(-x_p)\exp\left[\frac{-q(\phi_i - V_a)}{kT}\right] \tag{5.3.4}$$

结合这四个等式,可以用热平衡的值表示边界处过剩少数载流子浓度。定义过剩载流子浓度为

$$n' \equiv n - n_o \tag{5.3.5}$$

$$p' \equiv p - p_o \tag{5.3.6}$$

于是

$$n_p'(-x_p) = n_{po}(-x_p)\left[\exp\left(\frac{qV_a}{kT}\right) - 1\right] \tag{5.3.7}$$

$$p_n'(x_n) = p_{no}(x_n)\left[\exp\left(\frac{qV_a}{kT}\right) - 1\right] \tag{5.3.8}$$

等式(5.3.7)和(5.3.8)是非常重要的结论,将被用来求出pn结附近的准中性区少数载流子连续性方程的解。上述等式表明少数载流子浓度与外加电压呈指数关系,而假设多数载流子浓度对外加电压不敏感(一阶近似)。因为热平衡时少数载流子浓度的典型值比多数载流子浓度低11到12个数量级,等式(5.3.7)和(5.3.8)不会和小注入假设冲突,除非指数因子具有10^{10}或者10^{11}量级。在得到电流电压关系后,我们将考虑第二个假设(载流子准平衡)。

5.3.2 理想二极管分析

上面的讨论说明了我们为什么把注意力放在少数载流子上,并得到过剩少数载流子浓度与偏压的关系。在这个条件下,很容易求出准中性区连续性方程(等式(5.1.4))的解。我们首先在一系列理想条件下求解连续性方程,得到理想二极管分析。

首先考虑过剩空穴注入到n型区,这时体复合主要是通过产生-复合中心进行的。这样,

等式(5.1.4)中的(G_p-R_p)项可用等式(5.2.9)来描述。因为假设小注入，讨论过剩载流子寿命时得到的等式(5.2.15)也成立，所以5.1节讨论的准中性假设下的空穴连续性方程变为

$$\frac{\partial p_n}{\partial t}=D_p\frac{\partial^2 p_n}{\partial x^2}-\frac{p_n-p_{no}}{\tau_p} \tag{5.3.9}$$

式中下标n强调是n型区的空穴。

考虑实际中最简单的掺杂形式，即沿x方向均匀掺杂，同时考虑稳态($\partial p/\partial t=0$)，并将等式(5.3.6)中定义的过剩空穴浓度p'代入等式(5.3.9)，得到微分方程

$$0=D_p\frac{\mathrm{d}^2p_n'}{\mathrm{d}x^2}-\frac{p_n'}{\tau_p} \tag{5.3.10}$$

该方程具有简单的指数形式的解

$$p_n'(x)=A\exp\left(-\frac{x-x_n}{\sqrt{D_p\tau_p}}\right)+B\exp\left(\frac{x-x_n}{\sqrt{D_p\tau_p}}\right) \tag{5.3.11}$$

式中A和B是由边界条件决定的常数。等式(5.3.11)中，特征尺寸$\sqrt{D_p\tau_p}$被称为扩散长度，用L_p表示(p型区电子的扩散长度用L_n表示)。作为等式(5.3.11)的一个特殊应用，我们根据结到欧姆接触处n型区长度的W_B，如图5.5所示，考虑两个极端的情况。

长二极管 如果W_B比扩散长度L_p长，那么所有注入的空穴在穿越整个W_B之前就被复合掉了，这种二极管被称做长二极管。对于长二极管，L_p是注入的空穴复合前在中性区运动的平均距离(习题5.8)。因为p_n'必然会随x的增加而减小，所以等式(5.3.11)中的常数B必须等于零。解中的常数A则由等式(5.3.8)决定，$p_n'(x_n)$是外加电压的函数。这样完整的解是(如图5.6所示)

$$p_n'(x)=p_{no}(\mathrm{e}^{qV_a/kT}-1)\exp\left(-\frac{x-x_n}{L_p}\right) \tag{5.3.12}$$

利用等式(5.3.12)中过剩空穴浓度的表达式，可直接得到空穴电流的表达式。因为假设中性区电场可忽略，所以，空穴只有扩散电流，由等式(1.2.22)得

$$\begin{aligned}J_p(x)&=-qD_p\frac{\mathrm{d}p_n}{\mathrm{d}x}=qD_p\frac{p_{no}}{L_p}(\mathrm{e}^{qV_a/kT}-1)\exp\left(-\frac{x-x_n}{L_p}\right)\\&=qD_p\frac{n_i^2}{N_dL_p}(\mathrm{e}^{qV_a/kT}-1)\exp\left(-\frac{x-x_n}{L_p}\right)\end{aligned} \tag{5.3.13}$$

空穴电流在$x=x_n$处最大，远离结后逐渐减小，如图5.7所示，因为当载流子浓度因复合而减少时，空穴下降的梯度也减少。稳态时总的电流必须保持为常数，所以随着远离结的距离的增加，电子电流必须增加，提供与空穴复合的电子。

在欧姆接触W_B处总电流完全由电子电流组成，在朝结方向运动时，因为电子与注入的空穴复合导致电子电流减小。结区的电子电流是由n型区通过结注入到p型区的电子组成，注入到p型区的电子构成了p型区的少数载流子电流。所以，总电流等于这两种注入的少数载流子的电流之和：注入到n型区的空穴电流加上注入到p型区的电子电流。

注入到p型区的少数载流子电子电流，可以用与推导等式(5.3.13)类似的方法处理。如

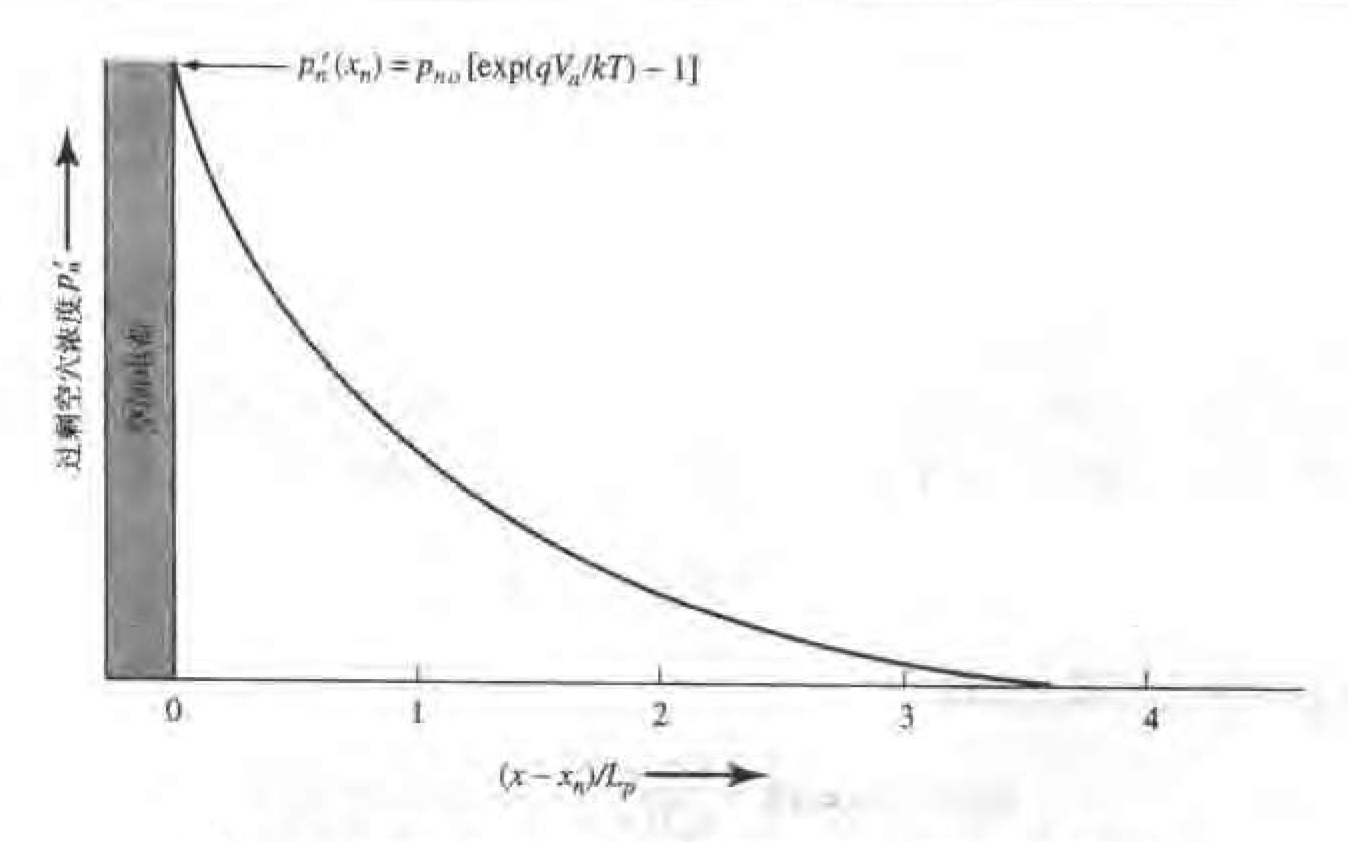


图 5.6 正向偏压 V_a 下长二极管准中性 n 型区的空穴浓度分布。过剩空穴浓度 p_n' 由等式(5.3.12)计算得到

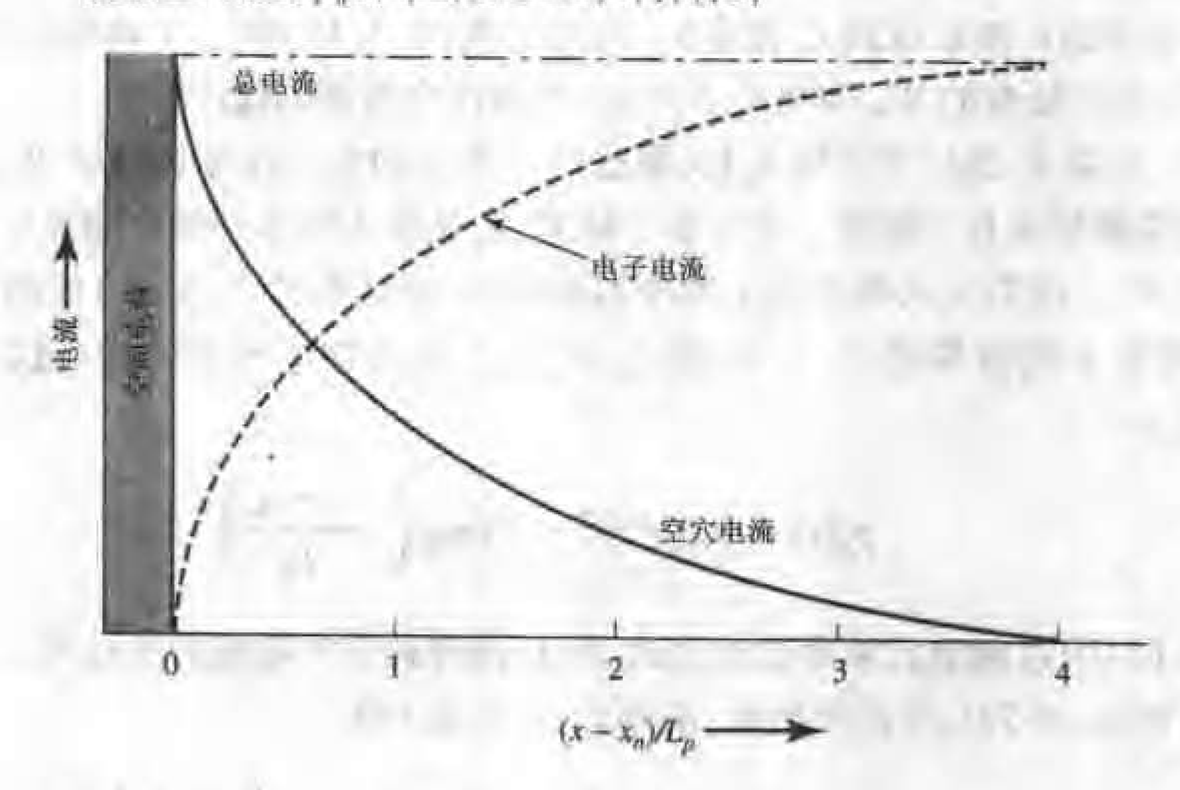


图 5.7 对于图 5.5 中的长二极管,准中性 n 型区的空穴电流(实线)和电子电流(虚线)分布。两种电流和 J(点划线)是常数。空穴电流由等式(5.3.13)计算得到

果欧姆接触是在 $-W_E$ 处,$W_E \gg L_n = \sqrt{D_n \tau_n}$,那么

$$J_n = qD_n \frac{n_i^2}{N_a L_n}(e^{qV_a/kT} - 1)\exp\left(\frac{x + x_p}{L_n}\right) \tag{5.3.14}$$

因为我们选择结的位置作为 x 轴的原点,x 在整个 p 型区为负。所以,与 n 型区的 J_p 一样,J_n 沿远离结方向减小。为了得到总电流 J_t 的表达式,我们利用等式(5.3.13)和(5.3.14)分别求出在 $-x_p$ 和 $+x_n$ 处少数载流子的电流分量,它们的和为

$$\begin{aligned} J_t = J_p(x_n) + J_n(-x_p) &= qn_i^2\left(\frac{D_p}{N_d L_p} + \frac{D_n}{N_a L_n}\right)(e^{qV_a/kT} - 1) \\ &= J_0(e^{qV_a/kT} - 1) \end{aligned} \tag{5.3.15}$$

当负偏压等于几个 kT/q 时，由这个等式计算出的电流趋向于 J_0，其中 J_0 为饱和电流密度。等式(5.3.15)与金属-半导体 Schottky 势垒二极管得到的等式(3.3.6)形式一样。相似的电流电压关系是因为两者都采用了准平衡假设，从而推导出等式(5.3.7)和(5.3.8)。

短二极管 另一个极端情形是 n 型和 p 型区长度 W_B 和 W_E 远小于少子扩散长度 L_p 和 L_n。在这种情况下，准中性区几乎没有复合发生，所有注入的少数载流子在二极管两端的欧姆接触处复合。在这种情形下，将等式(5.3.11)中的指数项用台劳级数展开，取前面两项得到

$$p_n'(x) = A' + B'\frac{(x - x_n)}{L_p} \tag{5.3.16}$$

因为 $x = W_B$ 处为欧姆接触，所以 $p_n'(W_B) = 0$。$x = x_n$ 处的边界条件由等式(5.3.8)给出，所以 n 型区过剩空穴浓度为

$$p_n'(x) = p_{no}(e^{qV_a/kT} - 1)\left(1 - \frac{x - x_n}{W_B'}\right) \tag{5.3.17}$$

式中 $W_B' = W_B - x_n$，是准中性 n 型区的长度，如图 5.5 所示。由等式(5.3.17)可看出，过剩空穴浓度在整个 n 型区随距离的增加线性减小，如图 5.8 所示。限制条件 $L_p \gg W_B$ 等价于所有的注入空穴在扩散通过整个 n 型区后才被复合。n 型区不发生复合的假设等价于等式(5.3.10)中的寿命 τ_p 接近无穷，因此微分方程有线性解。

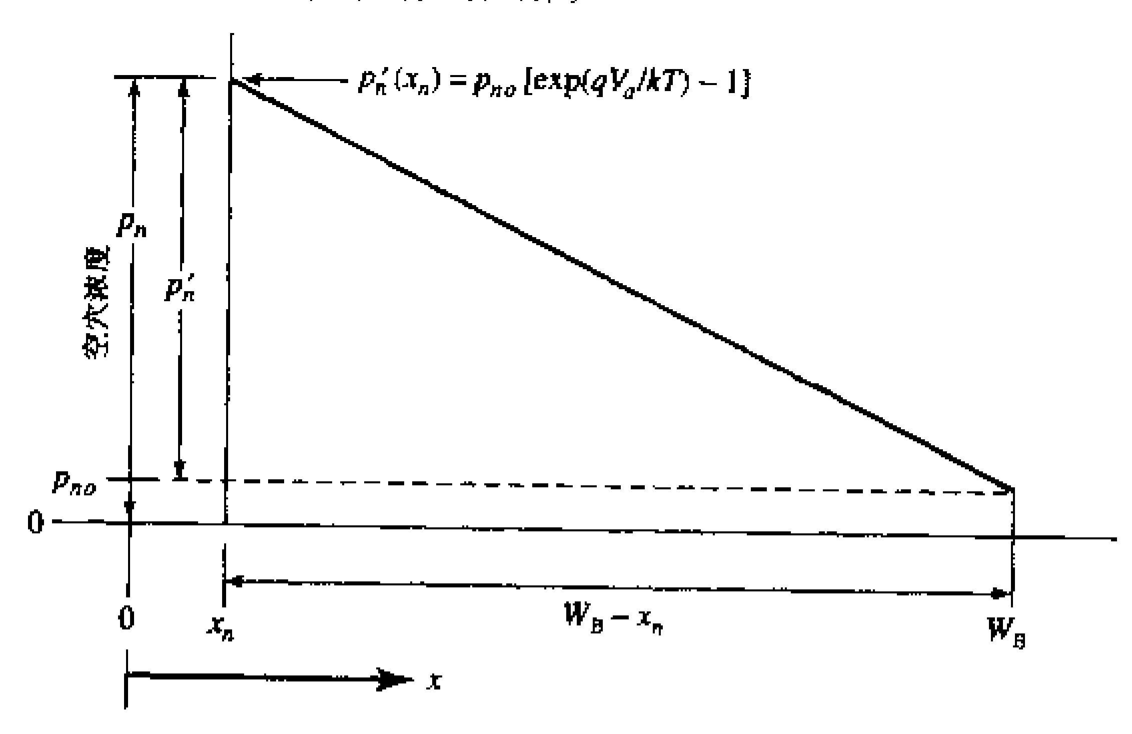


图 5.8 正向偏压 V_a 下理想的短二极管准中性 n 型区的空穴浓度分布。过剩空穴浓度 p_n' 由等式 5.3.17 计算得到

载流子浓度线性变化表明空穴电流在整个 n 型区保持为常数，不需要电子电流来补偿复合的空穴。因此

$$J_p = -qD_p\frac{dP}{dx} = qD_p\frac{p_{no}}{W_B'}(e^{qV_a/kT} - 1) = qD_p\frac{n_i^2}{N_d W_B'}(e^{qV_a/kT} - 1) \tag{5.3.18}$$

比较短二极管和长二极管的空穴电流表达式(5.3.18)和(5.3.13)，可看出二者很相似，只是与各自的几何结构相关的特征长度不同。长二极管的特征长度是少数载流子扩散长度，而短二极管的特征长度是准中性区的长度。与长二极管一样，短二极管中的总电流由注入到

p 型区的电子电流和注入到 n 型区的空穴电流组成。因此，对于 $W_E \ll L_n = \sqrt{D_n\tau_n}$ 和 $W_B \ll L_p = \sqrt{D_p\tau_p}$ 的短二极管，有

$$J_t = qn_i^2\left(\frac{D_p}{N_d W_B'} + \frac{D_n}{N_a W_E'}\right)(e^{qV_a/kT} - 1) \tag{5.3.19}$$

当然，给定的二极管可能是这两种极端情况的组合，也就是说，可能 p 型区长度远小于电子扩散长度，n 型区长度远大于空穴扩散长度，或者刚好相反。处理这样的情况也很容易。

两类二极管的电流电压关系(等式(5.3.15)和(5.3.19))都表明正向偏置下有大的电流流过，反偏时只有很小的饱和电流。从物理上看，这一不对称的现象是因为正向偏置促进了多数载流子从结的两侧注入，这些载流子在欧姆接触处被另一种载流子代替。而反向偏置下，通过结的净电流是由各区少数载流子组成，这些载流子数量较少，并在电极末端未被取代，因此，反偏时只有小的饱和电流流过(习题5.7)。

从等式(5.3.15)和(5.3.19)还可看出，反向饱和电流取决于结的轻掺杂一侧。例如，如果 n 型区的掺杂浓度远小于 p 型区，那么通过结注入到 n 型区的空穴电流就远大于注入到 p 型区的电子电流。这就是在轻掺杂的 n 型硅上扩散形成重掺杂 p 型区的情况。

例题　长二极管和短二极管

(a)在 700μm 厚，电阻率 $\rho = 4.5\Omega \cdot cm$ 的掺磷硅片上扩散高浓度的硼，形成硅二极管，并在硼扩散区和晶体的背面制作电极。pn 结(冶金结)距离表面 0.5μm，面积为 $10^{-4}cm^2$，可看做是单边突变结。如果衬底的空穴寿命 $\tau_{ps} = 1\mu s$，求二极管的饱和电流。

(b)另一个不同的工艺过程是在 3μm 厚，磷掺杂，电阻率 $\rho = 4.5\Omega \cdot cm$ 的外延层上以相似的扩散过程形成 p^+n 结。外延层淀积在重掺杂($N_d > 10^{19}cm^{-3}$)的 n 型衬底上。外延层的电子寿命 $\tau_{pe} = 1\mu s$，衬底的空穴寿命较短($\tau_{ps} = 100ps$)。pn 结距离表面 0.5μm，面积为 $10^{-4}cm^2$ 的，可看做是单边突变结。求二极管的饱和电流。

解：

(a)从图 1.15 或者表 4.1，可查出 $\rho = 4.5\Omega \cdot cm$ 的 n 型材料对应于 $N_d = 1 \times 10^{15}cm^{-3}$。从图 1.16 查出的扩散系数 $D_p = 12cm^2s^{-1}$。所以，扩散长度 $L_p = \sqrt{D_p\tau_p} = 34.6\mu m$。因为 L_p 远小于硅片厚度(700μm)，器件是长二极管，所以，饱和电流密度由等式(5.3.15)给出，饱和电流 $I_0 = J_0A = 1.2 \times 10^{-14}A$。

(b)本题的 pn 结外延层厚度远小于 L_p，几乎所有注入的空穴都能扩散到很容易发生复合的衬底上，所以这个结构是短二极管，J_0 由等式(5.3.18)给出。有效的中性 n 型区宽度是 1.5μm(外延层厚度减去结深，再减去大约 1μm 的扩展在轻掺杂 n 型区的耗尽区宽度)。利用等式(5.3.18)，计算得到二极管的饱和电流 $I_0 = J_0A = 2.7 \times 10^{-13}A$[7]。

从这个例题可看出，短二极管的饱和电流大，因为少数载流子浓度梯度增加了。

---

[7] 更精确地讲，短二极管的反向电流随着外加反偏电压的增加而增加，因为 pn 结耗尽层向轻掺杂外延层的中性区扩展，中性区宽度明显的减小增加了少数载流子浓度梯度和二极管的反向电流。

近似分析的适用范围 在进一步讨论pn结的物理特性前，我们再来考虑最初的两个假设。这些假设是：降落在准中性区的欧姆压降很小，这样V_a全部降落在空间电荷区上；外加偏压不会显著改变热平衡时的扩散和漂移之间的平衡。

为了分析这些假设，考虑一种典型的情形：$N_d=5\times10^{15}\text{cm}^{-3}$，$N_a=5\times10^{18}\text{cm}^{-3}$。先考虑短二极管，$W_B=W_E=3\mu\text{m}$，从等式(5.3.19)计算得到饱和电流密度大约是10^{-9}A cm^{-2}。横截面积的典型值为10^{-5}cm^2，正向偏压的典型值~0.65V，得到的电流大约是~1mA或者100A cm^{-2}。该电流基本上全部由注入到n型区的空穴电流组成，因为等式(5.3.19)中饱和电流的第一项远大于第二项。这样，欧姆电压仅降落在高电导率的p型区，该区域的电阻率是$0.03\Omega\cdot\text{cm}$，电场只有3V cm^{-1}。因此，$3\mu\text{m}$长度上的电压降近似为1mV，与外加电压0.65V相比可忽略。注意这种情况下基本上没有欧姆电压降落在n型区，因为该区的电子电流可忽略。

如果考虑两侧都是低掺杂的长二极管，该假设就可能不成立。为了说明这一点，假设$N_a=N_d=10^{16}\text{cm}^{-3}$，横截面为$10^{-5}\text{cm}^2$，外加电压为0.65V。当$L_n\simeq L_p=30\mu\text{m}$时，由等式5.3.15得出的电流大约是$32\times10^{-6}\text{A}$。如果p型区和n型区长度都是$100\mu\text{m}$，降落在p型区上的欧姆电压大约是0.03V，n型区上大约是0.02V。对于一阶近似，与外加电压0.65V相比，这两个电压可忽略。这些结果对很多实际情况是很典型的，因此可做出如下推断，对于低的或者中等大小的电流密度，假设整个外加电压只用于改变pn结的势垒高度是合理的。

如果外加正向电压接近于内建电势，那么阻挡多数载流子的势垒就被大大降低，将有大的电流流过。在这种情况下，相当大部分的外加电压降落在中性区，因此必须考虑中性区的串联电阻。利用电路分析，很容易将串联电阻效应考虑进来，这里不做详细考虑。

为了确定准平衡假设的适用范围，我们来比较热平衡下的漂移和扩散电流的典型值。典型的集成电路pn结的耗尽区宽度大约为10^{-5}cm，在这个范围内空穴浓度从10^{18}变化到10^4cm^{-3}，平均空穴扩散电流大约是10^5A cm^{-2}。那么由第4章可知，热平衡下这一空穴扩散电流正好等于耗尽区电场产生的方向相反的漂移电流。我们已经讨论过典型的正向偏置的二极管电流大约是10^2A cm^{-2}，仅为热平衡下两种电流的0.1%。这样，在处理小的或者中等大小偏压条件时，认为二极管只是略微偏离热平衡情形是合理的。这意味着利用等式(5.3.7)和(5.3.8)，就可以将结空间电荷区任意一侧的载流子浓度与结上的有效势垒高度(ϕ_i-V_a)联系起来。这两个等式的有效性还取决于小注入假设。热平衡下pn结上的相互抵消的扩散和漂移电流相对于实际流过的电流来说非常大，因此比等式(5.3.7)和(5.3.8)更普遍的关系式也是合理的。

根据等式(1.1.28)和(1.1.29)定义的准Fermi能级，可以推导出一个更普遍的关系式。准Fermi能级(ϕ_{fn}和ϕ_{fp})描述了非平衡条件下自由载流子浓度。热平衡下统一的Fermi能级决定了电子和空穴的浓度，但是在非平衡条件下，每一种载流子浓度由不同的准Fermi能级表示。

利用正向偏置pn结的讨论，可以预测在靠近外加偏置结的区域，由于注入明显改变了少数载流子的浓度，两个准Fermi能级分离。在远离结的区域，ϕ_{fn}和ϕ_{fp}统一。因为多数载流子所受影响可以忽略（迄今为止只讨论了低的或中等偏置条件），多数载流子的准Fermi能级基本上没有改变，而少数载流子的准Fermi能级需要修正[⑧]。所以外加偏压V_a等于

---

⑧ 另一种形象化描述多数载流子Fermi能级近似为常数的方法是利用等式(1.2.25)和(1.2.26)，其中用载流子浓度和准Fermi能级梯度的乘积表示总电流。对于给定的电流，载流子浓度越高，相应的Fermi能级梯度越低，因此，少数载流子的准Fermi能级的梯度比多数载流子的准Fermi能级梯度大很多。

$\phi_{fp}(x=-x_p)-\phi_{fn}(x=+x_n)$,即空间电荷区边界p型区一侧空穴和n型区一侧电子的准Fermi能级之差。将等式(1.1.30)应用于pn结空间电荷区($-x_p \leq x \leq +x_n$),得到

$$pn = n_i^2 \exp\left[\frac{q(\phi_p-\phi_n)}{kT}\right] = n_i^2 \exp\left(\frac{qV_a}{kT}\right) \tag{5.3.20}$$

等式(5.3.20)与由等式(5.3.3)和(5.3.4)得到的小注入情况下少数载流子浓度的边界值一致,在空间电荷区也是有效的[16]。当注入载流子浓度接近热平衡时的多数载流子浓度时(第7章),等式(5.3.20)也是有用的。

5.3.3 空间电荷区电流+

前面推导出的二极管方程(等式(5.3.15)和(5.3.19)),是在pn结的准中性区中分析得到的。空间电荷区只作为阻挡多数载流子扩散的势垒,它起到的惟一作用是建立了边界处的少数载流子浓度(等式(5.3.7)和(5.3.8))。一级近似下这是合理的,由它推导得到的方程(等式(5.3.15)和(5.3.19))称为理想二极管方程。在偏压范围较大时,理想二极管方程就不再精确,尤其是硅pn结,必须考虑pn结的空间电荷区对这些方程进行修正。

第4章指出,典型的空间电荷区宽度是10^{-4}cm。二极管的空间电荷区与准中性区一样,包含产生-复合中心。但是,这一区域的掺杂浓度变化很大,而且空穴和电子的浓度也变化迅速,这与准中性区不同。在正向偏压下,注入的载流子必须通过空间电荷区,部分载流子由于复合而消失。相反,在反向偏压下,空间电荷区中载流子的产生将形成电流,使得反向电流大于理想二极管方程所预示的饱和电流。

下面采用Shockley-Hall-Read理论推导空间电荷区的产生-复合电流表达式。为了简化计算,假定电子和空穴的俘获截面相等(等式(5.2.10)),外加偏压为V_a。pn乘积由等式(5.3.20)给出,这样总的复合率$U=-\mathrm{d}n/\mathrm{d}t=-\mathrm{d}p/\mathrm{d}t$就是

$$U = \frac{n_i^2(\mathrm{e}^{qV_a/kT}-1)}{\left(p+n+2n_i\cosh\dfrac{(E_t-E_i)}{kT}\right)\tau_0} \tag{5.3.21}$$

正向偏压下复合速率为正,反向偏压下为负(也就是说,反向偏压下,载流子净产生)。空间电荷区产生与复合形成的总电流由该区域复合率的积分给出

$$J_r = q\int_{-x_p}^{x_n} U\,\mathrm{d}x \tag{5.3.22}$$

尽管这个积分不容易求得,但可定性讨论电流电压关系。首先考虑正向偏置,由5.2节知,复合中心接近于禁带中间位置时最有效,如图5.3所示。这样,在等式(5.3.21)中取$E_t \approx E_i$,当$p+n$最小时,复合率达到最大值。再加上等式(5.3.20)给出的pn乘积的约束条件,可以很容易求出(习题5.15)U取最大值的条件

$$p = n = n_i \exp\left(\frac{qV_a}{2kT}\right) \tag{5.3.23}$$

对于典型的正向偏置条件,$p+n$之和远大于空间电荷区的n_i值。如果空间电荷区中U为最大值的范围为x',并对等式(5.3.22)积分起到主要作用,复合电流可以表示为

$$J_r = \frac{qx'n_i^2(e^{qV_a/kT}-1)}{2n_i(e^{qV_a/2kT}+1)\tau_0} \approx \frac{qx'n_i}{2\tau_0}\exp\left(\frac{qV_a}{2kT}\right) \tag{5.3.24}$$

式中 $\tau_0 = 1/N_t\sigma v_{th}$ 是复合中心密度为 N_t 的区域中过剩载流子的寿命。与准中性区载流子的复合电流不同，空间电荷区的复合电流随外加电压的变化关系是 $\exp(qV_a/2kT)$（假设 x' 与外加偏压无关）。这一不同的指数关系可以在实际的二极管中观测到，尤其是低电流时。在分析空间电荷区的复合时，因为像 τ_0 这样的项无法精确确定，所以也没有必要仔细分析 x'，实际上，经常用整个空间电荷区宽度 x_d 来近似 x'。利用这一假设，可得到正向偏压下理想二极管电流 J_i（等式(5.3.15)）和空间电荷区复合电流 J_r 的比值

$$\frac{J_i}{J_r} = \frac{2n_i}{x_d}\left[\frac{L_n}{N_a}+\frac{L_p}{N_d}\right]\exp\left(\frac{qV_a}{2kT}\right) \tag{5.3.25}$$

当偏压增加时，空间电荷区复合电流 J_r 与理想二极管电流相比不重要。同时，当缺陷浓度减小时，扩散长度增加，J_r 相对 J_i 的重要性减弱。以典型单边硅二极管为例，$L_n = 60\mu m$，$x_d = 0.25\mu m$，$N_a = 10^{16}cm^{-3}$，当 V_a 约大于 0.375V 时，J_i 就超过了 J_r。

在反向偏置下，方程(5.3.21)中的分子接近 $-n_i^2$，所以 U 为负值，表明空间电荷区有净的产生。根据等式(5.3.20)，pn 乘积变得非常小，在 p 和 n 相等时 U 取极大值，这时它们都小于 n_i。更细致的分析表明，在本征 Fermi 能级 E_i 与准 Fermi 能级的交点之间的空间电荷区 x_i 内，p 和 n 都远小于 n_i[7]。这个区域可能比总的空间电荷区宽度 x_d 小很多。在 x_i 以外的区域，p 或者 n 将大于 n_i，产生率快速下降。空间电荷区的净产生率可近似为最大产生率与宽度 x_i 的乘积

$$J_g = \frac{qn_ix_i}{2\tau_0} \tag{5.3.26}$$

这里仍假设最有效的复合中心位于 E_i。对于 p 型区重掺杂的单边 pn 结来说，几乎所有的空间电荷区延伸在轻掺杂 n 型半导体中，那么我们可求出总的空间电荷区宽度 x_d，以及具有最大产生率的区域 x_i

$$x_d = \left[\frac{2\epsilon_s}{qN_d}(\phi_i - V_a)\right]^{1/2} = \left[\frac{2\epsilon_s kT}{q^2N_d}\left(\ln\frac{N_dN_a}{n_i^2} - \frac{qV_a}{kT}\right)\right]^{1/2} \tag{5.3.27}$$

和

$$x_i = \left(\frac{2\epsilon_s kT}{q^2N_d}\right)^{1/2}\left[\left(\ln\frac{N_d}{n_i} - \frac{qV_a}{kT}\right)^{1/2} - \left(\ln\frac{N_d}{n_i}\right)^{1/2}\right] \tag{5.3.28}$$

式中 V_a 是外加偏压（反偏时 $V_a<0$）。反偏电压较大时，x_d 和 x_i 都决定于外加偏压的平方根，并且两者差别变小。图 5.9 给出了不同的施主浓度的单边突变结的 x_i/x_d 与外加电压的关系。因为实际二极管中复合中心的浓度随位置变化，通常不能精确确定，所以一般不用区分 x_i 和 x_d，并将 x_d 用于等式(5.3.26)。

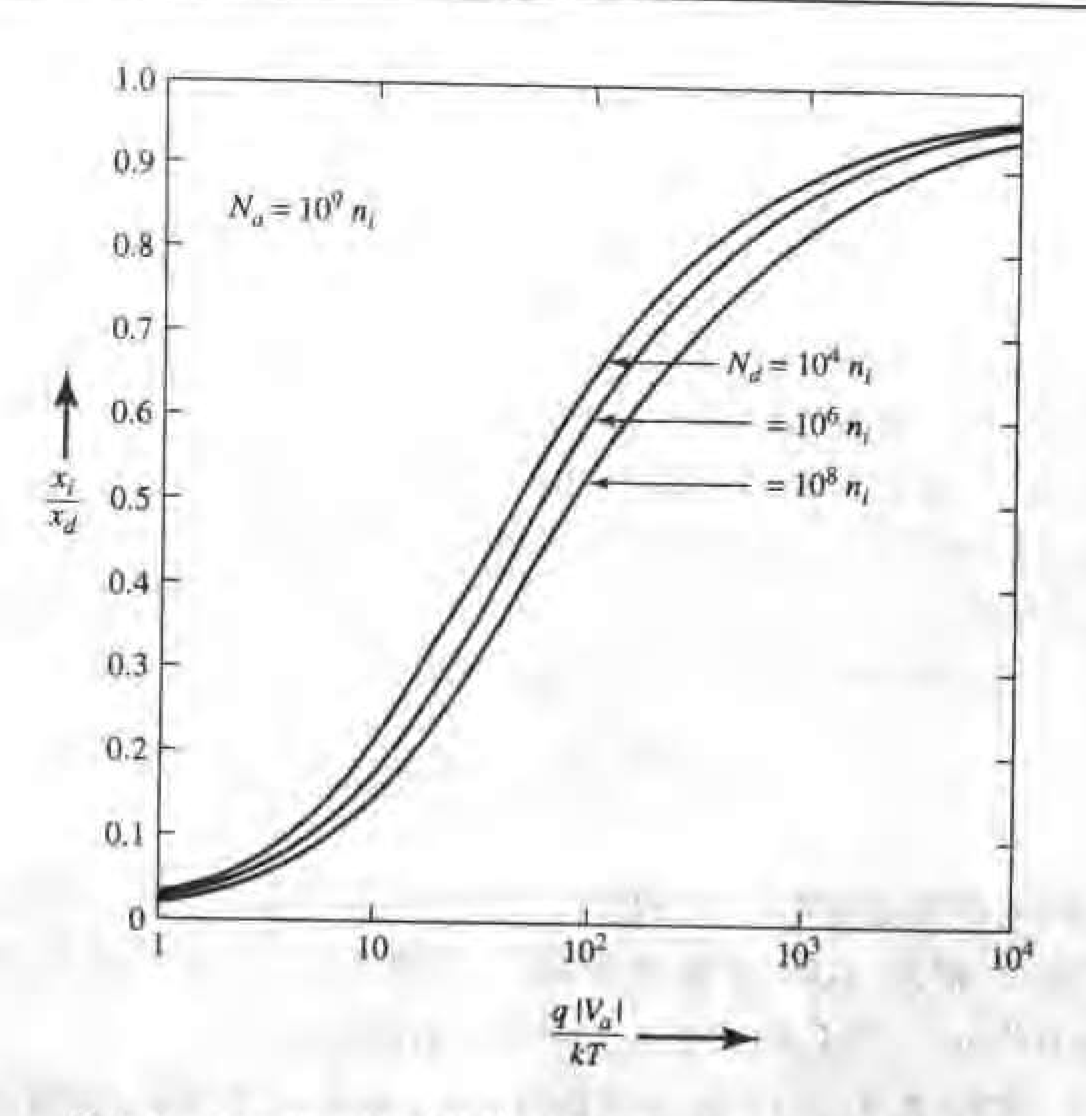


图 5.9 不同施主浓度的单边突变结，电荷产生区宽度 x_i 和与空间电荷区宽度 x_d 的比值与反偏电压的关系[7]

从等式(5.3.26)和(5.3.28)可看出空间电荷区的产生电流对反偏电压不是很敏感，基本上随反偏电压的平方根而变化。将等式(5.3.15)和等式(5.3.26)相比，可以估算出长二极管反偏电流中的准中性区电流与空间电荷区电流的相对大小

$$\frac{J_t}{J_g} = \frac{2n_i}{x_i}\left[\frac{L_n}{N_a} + \frac{L_p}{N_d}\right] \tag{5.3.29}$$

对于实际的硅二极管，比值 J_t/J_g 远小于1，表明反偏硅二极管的电流主要是在空间电荷区产生的，这也是造成反偏 Schottky 二极管(第3章)电流的原因之一。但是反偏 Schottky 二极管的大部分电流是与金属-半导体界面上的复合中心有关的。

例题 反偏二极管中的产生电流

(a)在 5Ω·cm 的 n 型厚硅片上扩散高掺杂硼形成硅二极管，在室温下比较反偏时的中性区和空间电荷区的产生电流。已知中性区和空间电荷区的寿命都是 10μs，外加偏压为 5V。

(b)比较 125℃时的这两种电流。假设寿命和扩散系数随温度的变化关系分别是 $T^{-1/2}$ 和 $T^{-2.2}$，如图 1.17 所示。

解：

(a)从图 1.15 或者表 4.1，得到 5Ω·cm 的 n 型硅的 $N_d = 9\times10^{14}\text{cm}^{-3}$。利用图 1.16，查出 $D_p = 12\text{cm}^2/\text{s}$。因此，300K 时的扩散长度 $L_p = \sqrt{D_p\tau_p} = 110\mu\text{m}$。对于 p^+n 结，电荷中性区的产生电流 J_t 为(等式(5.3.15))

$$J_t \approx qn_i^2 \frac{D_p}{N_d L_p} = 4.1 \times 10^{-11}\ \text{A cm}^{-2} = 41\ \text{pA cm}^{-2}$$

为了计算空间电荷区的产生电流，首先利用等式(5.3.28)确定产生区宽度 x_t。由已知条件得到 $x_t = 2.12\mu\text{m}$。由等式(5.3.26)计算出 $J_g = 24.6\text{nA cm}^{-2}$。比较这两个结果，可看出空间电荷区的产生电流大约是中性区产生电流的 600 倍。

(b)125℃时扩散系数 D_p 变为

$$D_p(398\text{K}) = D_p(300\text{K}) \times (300/398)^{2.2} = 12 \times 0.534 = 6.44\ \text{cm}^2\text{s}^{-1}$$

寿命也减少至 8.68μs。利用表 1.4 中的等式，计算得到 125℃时，$n_i = 6.5 \times 10^{12}\text{cm}^{-3}$。将这些值用于中性区得到，$L_p = 75\mu\text{m}$，$J_t \approx 6.45\mu\text{A cm}^{-2}$。因此，在 125℃时，两种电流的比值 $J_g/J_t = 2.1$。在更高的温度下两种电流几乎相等，这是因为一种电流随 n_i^2 增大，而另一种电流随 n_i 线性变化。

我们在讨论空间电荷区载流子的产生时，假设整个二极管区域的特性均匀。然而，pn 结二极管通常用绝缘的二氧化硅隔离。边界处硅晶格的不连续会造成浓度很高的产生-复合中心。实际上，空间电荷区与二极管边界隔离区相交处的载流子产生，通常是现代集成电路二极管漏电流的主要成分。如果器件工作要求漏电流必须很小，就需要非常仔细地控制工艺以减小空间电荷区漏电流的边界分量。这一表面(或者界面)产生电流将在第 8 章中详细讨论。

小结 关于 pn 结的电流分析篇幅较长，有必要总结一下这些内容。

求解两类均匀掺杂的突变结二极管的连续性方程，得到两类二极管的电流与电压关系(等式(5.3.15)和(5.3.19))，这与前面得到的 Schottky 势垒二极管类似，通常被称为理想二极管定律 $[J = J_0(\exp(qV_a/kT) - 1)]$。实际二极管的电流电压关系必须考虑空间电荷区的产生电流。当硅二极管在反向偏置或者低正偏时，这些电流将起主导作用。图 5.10 是长二极管在中等大小的正向偏压下的电流示意图。每一侧结附近的准中性区的总电流由以下几部分组成：(1)注入的少数载流子向远离结方向扩散，(2)向结区漂移的多数载流子与注入到准中性区的少数载流子的复合，(3)将被注入到掺杂类型相反的准中性区的向结方向漂移的多数载流子，(4)从两侧向结空间电荷区漂移的多数载流子在空间电荷区复合。每一侧远离结的区域，整个电流都是由漂移的多数载流子组成的。

下面分析 pn 结准中性区掺杂浓度不均匀的情况。在扩散结和注入结中会出现不均匀的掺杂分布，对这种情况最好采用将在 5.5 节讨论的数值方法分析，或者采用第 6 章中讨论晶体管的近似方法。我们将在第 6 章中看到，即使在实际的非均匀掺杂的情况下，理想二极管定律也成立，只是饱和电流 J_0 不同。对于非均匀掺杂二极管，空间电荷区的产生-复合将导致额外的电流成分。观察得到的稳态电流电压关系是理想二极管电流与这些额外分量的和。

5.3.4 异质结⁺

由前面的分析知，流过同质硅 pn 结的电流中，注入的电子和空穴之比与掺杂浓度成反比。为了减小向结的某一侧的电荷注入，该侧的掺杂浓度必须比另一侧高很多。如果掺杂浓度的

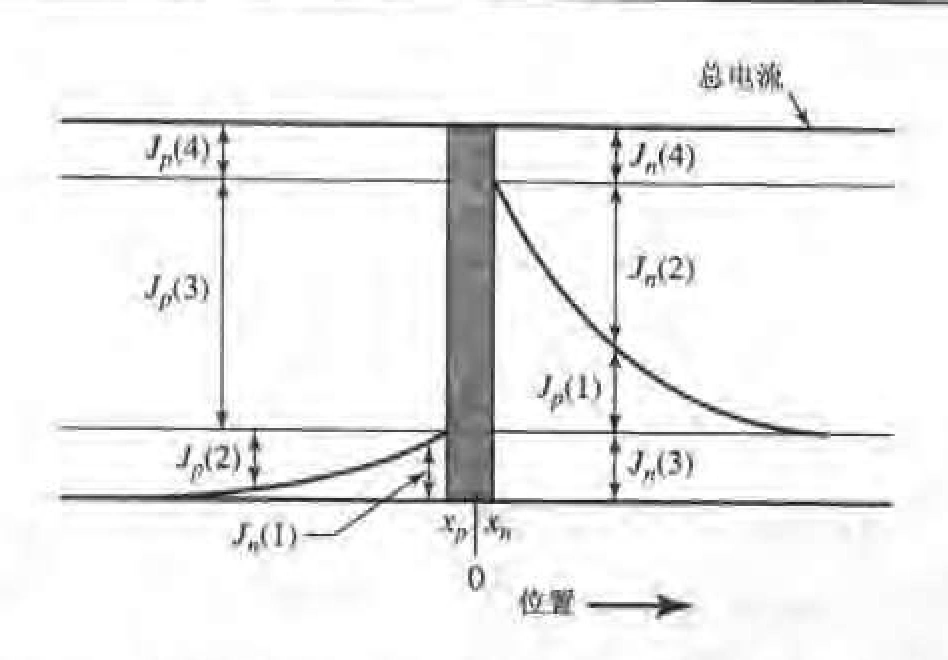


图 5.10 中等大小正向偏压下的长二极管准中性区电流各分量:$J(1)$为注入的少数载流子电流,$J(2)$为与$J(1)$复合的多数载流子电流,$J(3)$为注入通过结区的多数载流子电流,$J(4)$为空间电荷区复合电流

改变受到其他条件的制约,则用这种方法控制注入电流比就会受到限制,而注入电流比通常可以决定晶体管的增益。

然而,如果结两侧由不同电学特性的半导体材料组成,就可以用另一种方法控制注入电荷比。异质结的电子注入势垒 ϕ_{Bn} 和空穴注入势垒 ϕ_{Bp} 的不同,提供了另一个可控制二极管和晶体管特性的参数,我们将在第 6 章和第 7 章中讨论。

由第 4 章知,同质结的势垒 ϕ_{Bn} 和 ϕ_{Bp} 相等,都等于内建电势 ϕ_i;ϕ_i 由结两侧的功函数差决定,或者说是由半导体的带隙和掺杂决定。在异质结中,如图 4.9 和图 4.10 所示,势垒仍取决于功函数差,但还需要考虑两个带隙,所以势垒受到两个带隙以及掺杂浓度的影响,还可能受到结两侧导带边和价带边不连续的影响[8,9]。通常,$\phi_{Bn} \neq \phi_{Bp}$。因此,内建电势 ϕ_i 不是特别有用,我们主要考虑参数 ϕ_{Bn} 和 ϕ_{Bp}。

为了分析异质结电流,考虑两种情形:首先考虑电流与异质结空间电荷区中性边界处的少数载流子浓度及其分布有关,不考虑其他的势垒对电流的限制,如图 5.11(a)所示。这一传导机制对应于已经讨论过的同质结扩散电流(等式(5.3.13)和(5.3.18)),只是电子和空穴电流大小不同。然后考虑附加势垒对电流的限制,如图 5.11(c)所示。载流子热发射越过附加势垒(或者可能是隧穿通过势垒)将传导电流。

下面讨论第一种情况,如图 5.11(a)所示,扩散电流起主导作用,导带尖峰(由 n 型材料中 E_c 的弯曲决定)没有限制载流子通过结的输运。在这种情况下,由 4.2 节知,势垒由两种半导体中性区的性质决定(即热平衡时整个系统的 Fermi 能级必须一致)。对于电子

$$\phi_{Bn} = \frac{E_{c1} - E_{c2}}{q} \tag{5.3.30}$$

对于空穴

$$\phi_{Bp} = \frac{E_{c1} - E_{c2} + \Delta E_g}{q} \tag{5.3.31}$$

式中 E_{c1} 和 E_{c2} 分别是材料 1 和材料 2 的中性区导带边的能量。

考虑两种半导体材料中的能带弯曲的代数和(内建电势 ϕ_i),以及界面处导带和价带边的

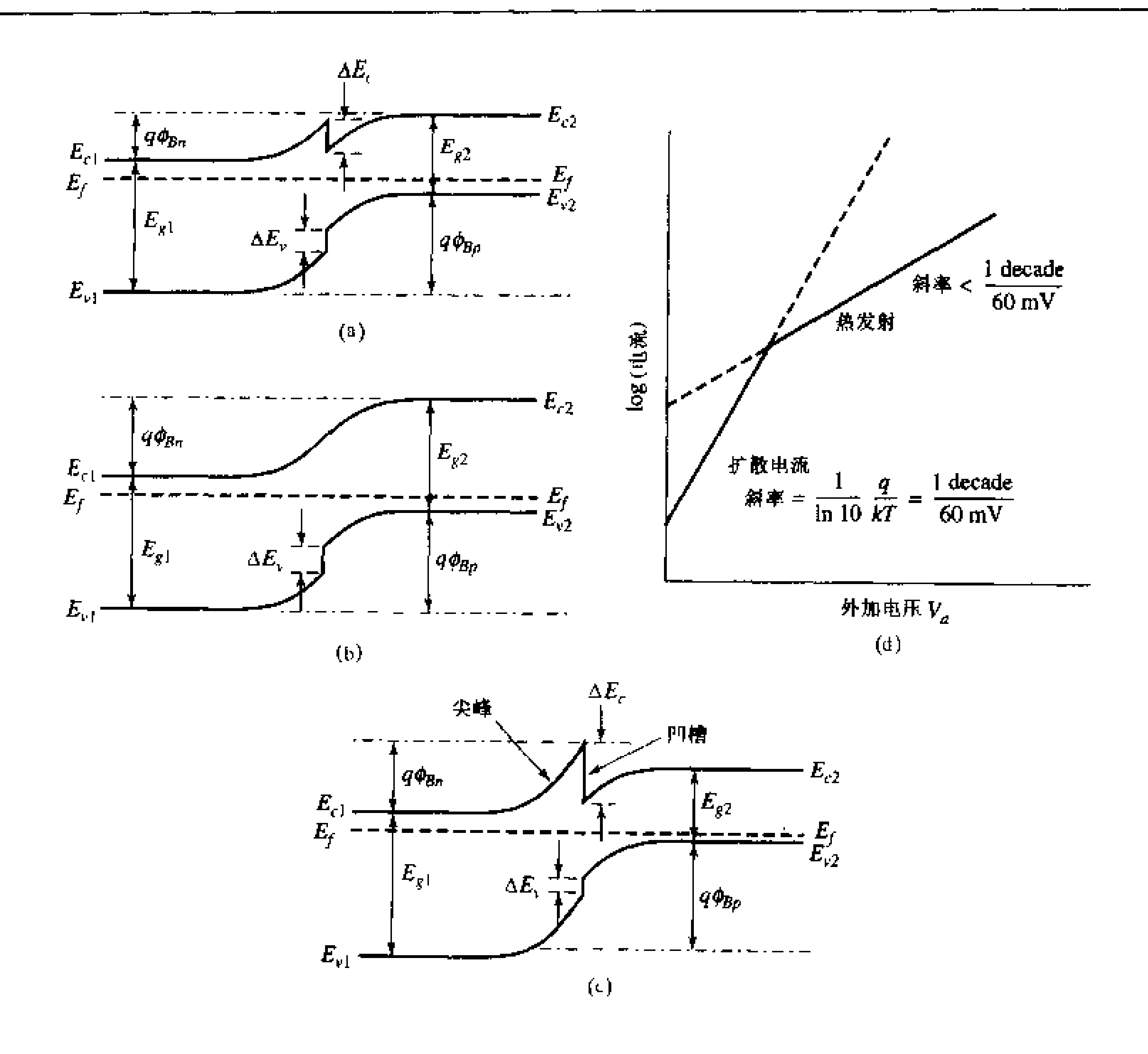


图5.11 (a)两侧带隙与电子亲和势X不同的异质结导带和价带边 材料1中的导带尖峰比材料2的导带边低;(b)$X_1 = X_2$时的能带图;(c)$X_1 \neq X_2$,且材料1中的导带尖峰比材料2的导带边高;(d)(引自[8])边界条件介于(a)和(c)之间的二极管的电流电压特性,给出了低电压下的扩散电流和高偏置下的热电子发射电流

不连续ΔE_c和ΔE_v,我们还可以用内建电势来表示势垒。这样,$\phi_{Bn} = \phi_i - \Delta E_c/q$,$\phi_{Bp} = \phi_i + \Delta E_v/q$。然而,这些表达式可能会误导大家,因为$\phi_i$本身依赖于$\Delta E_c$。尽管势垒看上去只依赖于价带边或者导带边的不连续性,因为等式(4.2.20)表明内建电势随带边不连续而增加,所以$\phi_{Bn} = \phi_i - \Delta E_c/q$与$\Delta E_c$无关,而$\phi_{Bp} = \phi_i + \Delta E_c/q$则取决于整个带隙不连续值$\Delta E_g$。

为了得到扩散电流的表达式,需要将中性区的少数载流子浓度和材料性质联系起来,尤其是和两种半导体的带隙联系起来。与4.2节相同,我们只考虑一种非常重要的情况,n型材料1具有宽带隙E_{g1},p型材料具有较小的带隙E_{g2},如图5.11(a)所示。我们可以用势垒高度,或者直接根据材料性质和掺杂浓度来表示结两侧的载流子浓度。

首先考虑结上没有外加电压时的电子浓度。n型一侧材料1的中性区多数载流子电子浓度是

$$N_d = n_{no1} = N_{c1}\exp\left(-\frac{E_{c1} - E_f}{kT}\right) \tag{5.3.32}$$

p 型一侧材料 2 中的中性区少数载流子电子浓度是

$$n_{po2} = N_{c2}\exp\left(-\frac{E_{c2}-E_f}{kT}\right) \tag{5.3.33}$$

通过结的多数和少数载流子浓度和势垒高度 ϕ_{Bn} 有关(等式 5.3.30)。

利用少数载流子浓度的表达式来求正向偏压 V_a 下通过 pn 异质结的电流。假设半导体中性区只有小的电压降落,这样势垒高度 ϕ_{Bn} 和 ϕ_{Bp} 由于外加偏压 V_a 的作用将分别减小。中性区边界上的少数载流子浓度与同质结(等式(5.3.3))相同,增加了 $\exp(qV_a/kT)$。根据 p 型半导体中性区边界的电子浓度,得到长二极管中 p 型区少数载流子电子电流(等式(5.3.14))是

$$J_n = qD_n\frac{dn_{p2}}{dx} = -\frac{qD_n}{L_n}n_{po2}\left[\exp\left(\frac{qV_a}{kT}\right)-1\right] = -\frac{qD_n}{L_n}\frac{n_{i2}^2}{N_{a2}}\left[\exp\left(\frac{qV_a}{kT}\right)-1\right] \tag{5.3.34}$$

这样,注入电子电流与由窄带隙材料 2 构成的同质结情况一样。在讨论异质结双极晶体管之前,我们希望用材料 1 中的本征载流子浓度来表示电流

$$J_n = -\frac{qD_n}{L_n}\frac{n_{i1}^2}{N_{a2}}\exp\left(\frac{\Delta E_g}{kT}\right)\left[\exp\left(\frac{qV_a}{kT}\right)-1\right] \tag{5.3.35}$$

这样强调了流过异质结的电子电流比由材料 1 构成的同质结要大很多,因为阻止电子注入的势垒高度 ϕ_{Bn} 降低了。对于短二极管,为了得到相似的表达式,只需用 p 型准中性区的长度 W_p' 取代 L_n。

对空穴电流同样用材料 1 的本征载流子浓度写下类似的电流表达式

$$J_p = -qD_p\frac{dp_{n1}}{dx} = -\frac{qD_p}{L_p}p_{no1}\left[\exp\left(\frac{qV_a}{kT}\right)-1\right] = -\frac{qD_p}{L_p}\frac{n_{i1}^2}{N_{d1}}\left[\exp\left(\frac{qV_a}{kT}\right)-1\right] \tag{5.3.36}$$

等式(5.3.36)表明从材料 2 注入到材料 1 的空穴电流与由材料 1 构成的同质结相同,因为同质结和异质结的势垒高度 ϕ_{Bp} 是一致的。就像预测的那样,等式(5.3.35)和(5.3.36)与导带不连续无关,仅与两种材料的带隙和掺杂有关。同样,我们可以用电子和空穴的势垒高度 ϕ_{Bn} 和 ϕ_{Bp} 来表示通过结注入的电流。

我们可以举例说明电流与不同的电子和空穴的注入势垒的关系。考虑以下异质结,一侧是 n 型硅,另一侧是由硅和锗($Si_{1-x}Ge_x$)合金组成的 p 型半导体,这里 x 代表合金中锗原子占的比例。$Si_{1-x}Ge_x$ 合金的带隙比硅小,电子亲和势与硅相似,这样导带的不连续很小,可以忽略,如图 5.11(b)所示,几乎所有的带隙差表现在价带边的不连续上。

当两种材料密切接触时,电荷迁移使得结两侧的 Fermi 能级一致,如图 5.11(b)所示。孤立的 $Si_{1-x}Ge_x$ 的 Fermi 能级与相同掺杂的 Si 更接近真空能级,因为它们的电子亲和势相同,且 $Si_{1-x}Ge_x$ 带隙窄。这样当两块半导体材料密切接触时,异质结中 Fermi 能级的移动要比相同掺杂的同质结小,因而阻挡电子从 n 型 Si 向 p 型 $Si_{1-x}Ge_x$ 运动的势垒比类似的同质结低。然而,空穴的势垒与同质结的势垒是相同的;价带的不连续(根据定义)补偿了 $Si_{1-x}Ge_x$ 功函数的减少。

外加正向偏置时,载流子注入遇到的结的势垒减小,p 型材料的电子浓度增加,有电流流过。异质结比同质结阻挡电子的势垒低,使得电子电流的注入更容易,因而开启或者阈值电压比类似掺杂的硅 pn 结二极管要低。然而空穴电流是与同质结相似的。这样,电子和空穴电流

的比值增加，这一点非常重要。反过来看，要得到给定的电子电流，异质结需要的偏压低，此时，注入到n型区的空穴电流比同质结低。

例题　异质结电流

考虑重掺杂n型硅和中等掺杂的p型$Si_{1-x}Ge_x$组成的异质结。假设，$Si_{1-x}Ge_x$的带隙随锗含量线性变化（不完全正确）。在半对数坐标系中，画出$x=0$、0.1、0.2和0.3（合金中锗含量是0、10%、20%和30%）的电流。假设外加正偏压为1V时，硅同质结电流是1mA。

解：　对于图5.11(b)所示的能带结构，扩散电流起主导作用。异质结中宽带隙半导体是重掺杂的n型硅，电流的主要部分是电子电流。所以利用等式(5.3.35)求出的电流表明，电流随带隙差指数递增

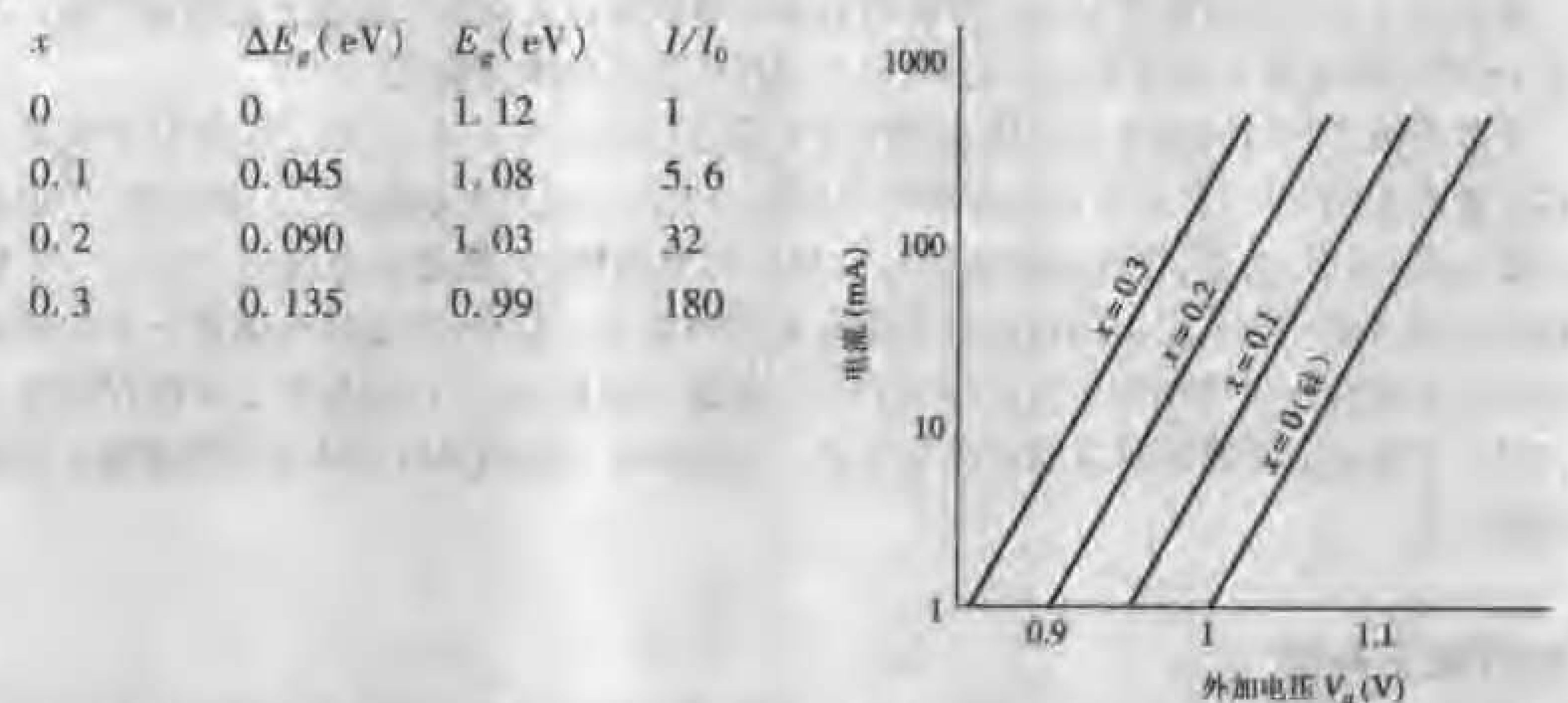
$$\frac{I}{I_0}=\exp\left(\frac{\Delta E_g}{kT}\right)$$

随着锗含量增加，$Si_{1-x}Ge_x$合金的带隙减小：

$$E_g=1.12-x(1.12-0.67)(\text{eV})=1.12-0.45x(\text{eV})$$



那么，



x	ΔE_g(eV)	E_g(eV)	I/I_0
0	0	1.12	1
0.1	0.045	1.08	5.6
0.2	0.090	1.03	32
0.3	0.135	0.99	180

注意，异质结二极管中的外加电压减小$\Delta E_g/q$时，电流与硅同质结二极管的相同。

现在我们讨论其他限制电流的因素。图5.11(b)所示的只是最优化的情况，外加偏压有效地降低了势垒高度，使得载流子可通过pn结。另外，$\Delta E_c=0$避免了电子在凹槽中被限制或散射的问题。当ΔE_c增加，或者p型区的掺杂增加时，情况将变得复杂（如果p型区是晶体管的基极，则需高掺杂以减小串联电阻，这将在第6和第7章中涉及）。当ΔE_c增加时，电子限制在凹槽中的情况变得更重要。

更严重的是，当导带边的尖峰变得比材料2中性区导带边高（见图5.11(c)）时，以上分析将失效。电子从异质结n型一侧向结的迁移，受到热电子发射的限制，只有能量超过尖峰势垒的电子才能穿越，$J_n\propto\exp(-q\phi_{Bn}/kT)$，这里的$\phi_{Bn}$比$(E_{c2}-E_{c1})/q$要大。当外加正向偏压时，每一侧的半导体的空间电荷区上的压降只是总电压的一部分。只有降落在材料1上的外加电压部分减小了电子从n型材料1向p型材料2运动时遇到的势垒ϕ_{Bn}时，电流才会增加，并且

电流随电压的增加比 $\exp(qV_a/kT)$ 慢。

尖峰是否限制电流取决于 n 型区和 p 型区的掺杂以及结上的偏压。由 4.2 节知，当 p 型区重掺杂时，尖峰的限制作用将变大，因为大部分的能带弯曲发生在 n 型区（等式(4.2.30)和(4.2.31)）。即使在低的正向偏压下尖峰没有限制电流流动，当偏压增加时，由于结两侧能带弯曲的相对改变，也有可能导致尖峰限制电流。图 5.11(d)（引自文献[8]中的图 1.4）给出了半对数坐标系下的低正向偏置下的扩散电流和高电压下的热电子发射电流与电压关系。实际上，采用更为复杂的工艺可以减小尖峰和凹槽，例如，在两种材料之间采用组分缓变的材料结构

由上述关于异质结电流的简要讨论中，我们看到异质结提供了额外的参数，来控制正向偏置下通过结的电子和空穴电流的注入比。这种控制对于设计异质结双极晶体管来说非常有用，我们将在第 6 和第 7 章中讨论。我们将在 7.6 节中看到空穴电流的减少可以增加异质结双极晶体管的高频性能，使之远好于同质结双极晶体管。

5.4 电荷存储与二极管瞬变特性

上一节我们看到通过 pn 结的正向偏压将导致电子从 n 型区向 p 型区注入，空穴则方向相反。这些注入的少数载流子在准中性区的分布将导致电流流动和二极管上的电荷存储。本节讨论 pn 结存储电荷及其与电流的关系，以及对外加偏压的瞬变响应。

少数载流子与时间的关系可用连续性方程描述（等式(5.1.4)），这一偏微分方程是与时间和位置有关的，可以针对不同的作用力函数和初始条件得到少数载流子的瞬态解，但通常不这么做，原因如下：首先，仅在特殊情况和理想作用力函数（只是真实电路的近似）下，才能得到精确的解析解，所以，求出的精确解并没有实际的意义。其次，不直接求偏微分方程是因为二极管瞬变特性不仅受到准中性区存储的少数载流子的控制，而且还与耗尽区的存储电荷有关，这样，必须同时计算两种存储电荷的变化。处理瞬态问题的最好方法是直接考虑这些电荷的物理行为。

少数载流子存储

存储在单位面积下准中性 n 型区注入的少数载流子总电荷，可以通过对整个准中性区的过剩空穴分布积分得到

$$Q_p = q\int_{x_n}^{W_n} p_n'(x)\,\mathrm{d}x \tag{5.4.1}$$

因为准中性条件，该区域电子将比热平衡时增加 Q_p/q（单位面积），以平衡空穴电荷 $+Q_p$。

首先考虑长二极管，所有注入的少数载流子在达到 n 型区电极前就复合了。为了简化，假设半导体均匀掺杂。由等式(5.3.12)知，随着空穴向 n 型区的进一步扩散，少数载流子浓度指数衰减。将等式(5.3.12)代入等式(5.4.1)，并在整个准中性区从空间电荷区边界开始积分，得到存储在 n 型区的空穴电荷 Q_p

$$Q_p = qL_p p_{no}(\mathrm{e}^{qV_a/kT} - 1) \tag{5.4.2}$$

存储的电荷与电流一样，对外加偏压呈指数依赖关系。无论是对于电流还是存储电荷，指数项来源于空间电荷区边界上的过剩少数载流子浓度。用相应的空穴电流来表示存储电荷，可

以加深物理理解 在等式(5.4.2)中利用 $x=x_n$ 处空穴电流(等式(5.3.13)),得到简单的表达式

$$Q_p = \frac{L_p^2}{D_p}J_p(x_n) = \tau_p J_p(x_n) \tag{5.4.3}$$

这样,对于这种理想的长二极管,存储电荷等于电流和寿命的乘积。这一关系是合理的,因为注入的载流子寿命越长,则向 n 型区扩散得越远,那么更多的空穴将被存储。

理想的短二极管的解略有不同。将空穴分布表达式(方程(5.3.17))代入等式(5.4.1),得到存储的空穴电荷是

$$Q_p = \frac{q(W_B - x_n)}{2}p_{no}(e^{qV_a/kT} - 1) \tag{5.4.4}$$

其中 W_B 是 n 型区宽度(图 5.5)。同样也可以用空穴电流(方程(5.3.18))来表达存储的空穴电荷

$$Q_p = \frac{(W_B - x_n)^2}{2D_p}J_p \tag{5.4.5}$$

尽管等式(5.4.5)中 J_p 前的分式具有时间的量纲,但并不是类似等式(5.4.3)中的寿命。为了得到这组常数的物理意义,我们将等式(5.4.5)重写为

$$\frac{Q_p}{J_p} = \frac{(W_B - x_n)^2}{2D_p} \tag{5.4.6}$$

存储的电荷除以电荷进入或离开 n 型区的速率正好等于载流子平均停留在这一区域的时间。所以,等式(5.4.6)的右侧等于短二极管中一个空穴穿越 n 型区的平均渡越时间 τ_{tr}。

少数载流子存储的瞬变特性[†] 我们可以通过少数载流子存储的物理图像来定性讨论 Q_p 的瞬态积累和衰减。考虑初始状态没有电荷存储的理想长二极管,突然外加正的恒定电流源,分析 n 型准中性区空穴的变化。在达到稳态前,空穴必须传输到准中性的 n 型区以建立起特定的空穴分布。理论上,在 $t=0^+$ 时,因为 n 型区几乎没有空穴,所以空穴浓度梯度很大。实际的空穴浓度梯度是由电流源决定的。存储的空穴电荷 Q_p 随着时间的增长,结上的电压增加反映出空穴浓度的边界值。在这种情形下,空穴的瞬态增加如图 5.12 所示。达到稳定的空穴分布所需的时间由稳态的空穴存储电荷与电流源的比值给出(忽略空穴的复合以及空间电荷区边界附近的界面陷阱存储的电荷)。

二极管关断时间受限于从准中性区移出存储空穴的速度。当反向偏压突然加在正向偏置的结上时,电流可以很快地改变方向,因为空间电荷区边界附近存储的空穴的微小变化就能改变梯度方向,如图 5.13(a)所示的曲线 1。只要空间电荷区边界能提供足够的少数载流子,二极管就可以在反向时传导大电流,结将继续保持正向偏置,直到耗尽区边界附近注入的少数载流子全部被移走。这意味着电流对时间的曲线(见图 5.13(b))最初几乎是常数(由外部电路决定),一直保持到 t_4,这时 $p_n'(x_n)$ 变为 0。随后,如图 5.13(a)所示的曲线 5 和 6,通过结传输空穴的速率变得越来越小,这时电流的衰减取决于准中性区的少数载流子寿命。

因为二极管的开关时间取决于必须被注入或者移走的存储电荷量,所以可通过减少存储电荷来缩短开关时间。从等式(5.4.3)可看出,存储电荷的减少可以通过限制正向电流,或者减小少数载流子寿命来实现。因此,优化设计二极管,不一定需要采用尽可能纯的材料来增大少数载流子寿命。实际上,有些应用要求器件的开关速度必须足够快,需要在半导体中有意引

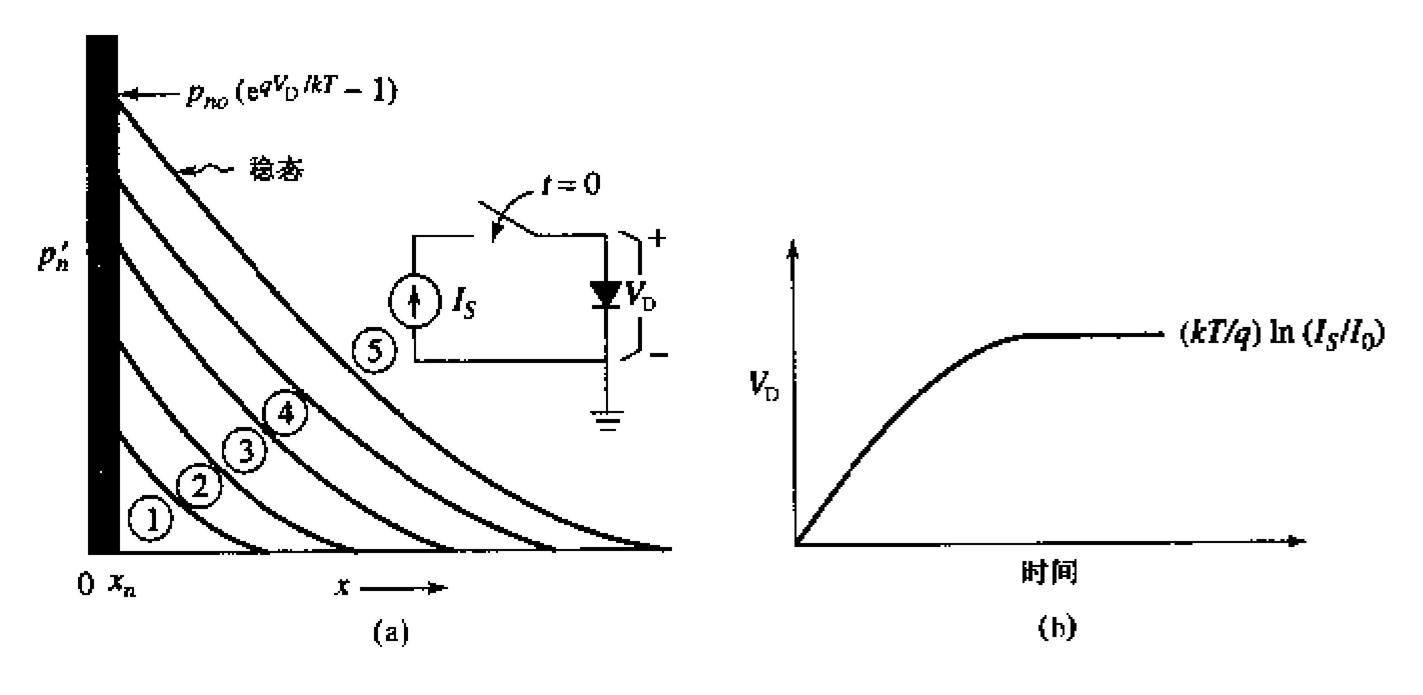


图 5.12 (a)理想长二极管初始没有偏置,在时间为零时突然施加恒流源驱动,存储的过剩空穴随时间变化。注意从时间(1)到时间(5),$x=x_n$处的梯度不变,表明注入的空穴电流为常数(电路显示在小图中);(b)二极管电压 V_D随时间的变化曲线

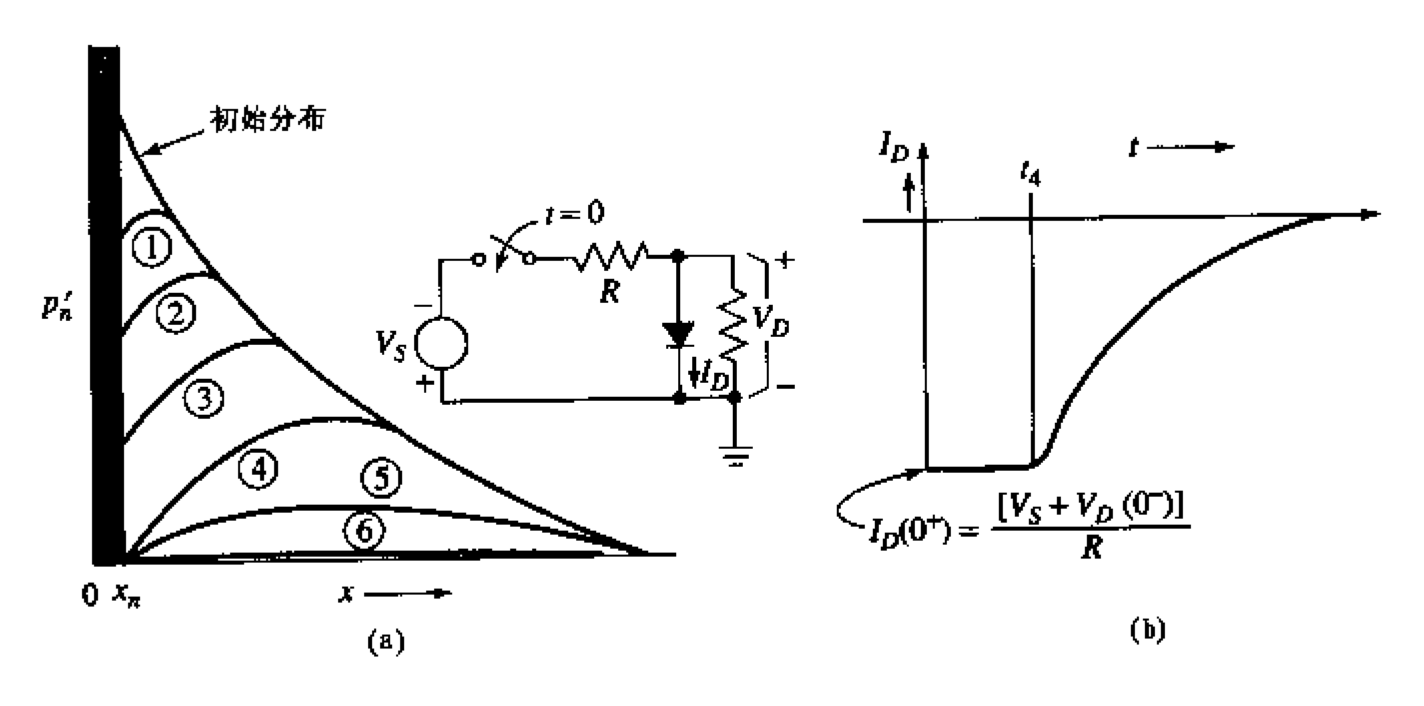


图 5.13 (a)理想长二极管中存储的过剩空穴的瞬态衰减过程。图中表示最初通过串联电阻施加正向偏置,在时间 $t=0$ 时突然改变成负偏压(电路在小图中显示);(b)二极管电流 I_D随时间的变化曲线

入复合中心以减小寿命,这样只有少量的电荷被存储。当然,关键是以可控的方式引入这种"减小寿命"的杂质,这样存储电荷的数量减小的同时,不会使反向漏电流增大到不可接受的水平(等式(5.3.15)中的 $-J_0$)。例如,在精确控制的温度下,在半导体硅片中掺金。

限制正向电流是缩短开关时间的另一种方式。集成电路中通常采用这种方式,以避免可能由掺杂引起的相邻器件的半导体材料性能的退化。流过 pn 结的正向电流可以采用 3.6 节讨论过的 Schottky 二极管箝位技术来限制。因为 Schottky 二极管比普通的 pn 结具有更低的开启电压,可通过在 pn 结上并联 Schottky 二极管来限制其电压和电流,这样 pn 结中的存储电荷就减少了,关断时间也缩短了。虽然会有部分电流流过 Schottky 二极管,但是只有非常少量的少数载流子电荷被存储。

迄今为止,我们已经讨论了 pn 结中 n 型一侧准中性区的空穴存储。当然,p 型区也有类似的电子存储,可以推导出类似于等式(5.4.2)和(5.4.4)的表达式。存储在二极管中的总的

少数载流子电荷是这两部分之和。

我们将会发现少数载流子的电荷存储、渡越时间以及它们之间的相互关系在讨论pn结瞬变特性时非常有用。处理瞬态问题可以忽略多数载流子浓度的改变，因为多数载流子在介电弛豫时间 $\tau_r = \epsilon_s/\sigma$ 内就能响应电场的改变，而电阻率为0.1Ω·cm的硅的介电弛豫时间为 10^{-13}s量级(0.1ps)。该时间远小于少数载流子的寿命(等式(5.4.3)和等式(5.4.6))。集成电路中的少数载流子寿命以及渡越时间的典型值介于 10^{-12}s到 10^{-3}s之间。

总的存储效应　4.3节讨论了pn结存储电荷的一种方式。当耗尽区宽度随反偏电压的改变而变化时，耗尽区边缘附近的多数载流子也在移动。这一耗尽区的电荷存储效应可以用小信号电容来模拟。由耗尽区电荷存储而产生的电容通常用 C_j 表示。由等式(4.3.8)表示的突变结耗尽电容就是一个例子。

类似的，在正向偏置条件下，准中性区存储的少数载流子电荷的变化也可以用另一种小信号电容来模拟。这个电容一般被称为扩散电容(定义为 C_d)，因为少数载流子是通过扩散通过准中性区的。根据定义 $C_d = \mathrm{d}Q_p/\mathrm{d}V_a$，利用长二极管的等式(5.4.2)，或者短二极管的等式(5.4.4)可求n型区存储的空穴对 C_d 的贡献(因为 Q_p 代表单位面积电荷，所以 C_d 代表单位面积电容)。因为这两个二极管具有同样的电压关系，所以一般情况下

$$C_d = \frac{qQ_{po}e^{qV_a/kT}}{kT} \tag{5.4.7}$$

其中对长二极管($J_{p0} = qD_p p_{no}/L_p$)

$$Q_{po} = J_{p0}\tau_p = qp_{no}L_p \tag{5.4.8}$$

对短二极管($J_{p0} = qD_p p_{no}/(W_B - x_n)$)

$$Q_{po} = J_{p0}\tau_{tr} = qp_{no}(W_B - x_n)/2 \tag{5.4.9}$$

同样，如果p型区的电子存储很重要的话，可在 C_d 中直接加入该分量。正如等式(5.4.7)指出的，C_d 在反偏时可被忽略，因为少数载流子的存储非常少。正向偏置下的 Q_p 和 C_d 都随电压的增加而指数增加。

从前面的讨论可看出，空间电荷区的电荷存储(用 C_j 代表)与准中性区的电荷存储的相对重要性强烈依赖于结电压。反向偏置下，准中性区的存储可以忽略，结电容的存储效应占主导地位。正向偏置下，尽管 C_j 增加(因为 x_d 减小了)，但 C_d 等式中的指数因子使得扩散电容以及和它相关的电荷存储占主导地位。然而，对于需要精确求解的实际问题，有必要计算正向偏置下两种类型的电荷存储。

例题　结和自由载流子存储

已知突变长二极管，$L_p = 20\mu$m，p型区掺杂浓度 $N_a = 10^{19}$cm^{-3}，n型区掺杂浓度 $N_d = 10^{16}$cm^{-3}。画出电荷 Q_p 关于外加电压 V_a 的变化曲线，其中 $-3 < V_a < 0.6$V。在同样的坐标轴上再绘制出当结偏置时，n型区一侧耗尽区宽度改变对应的存储电荷的增加 Q_T。

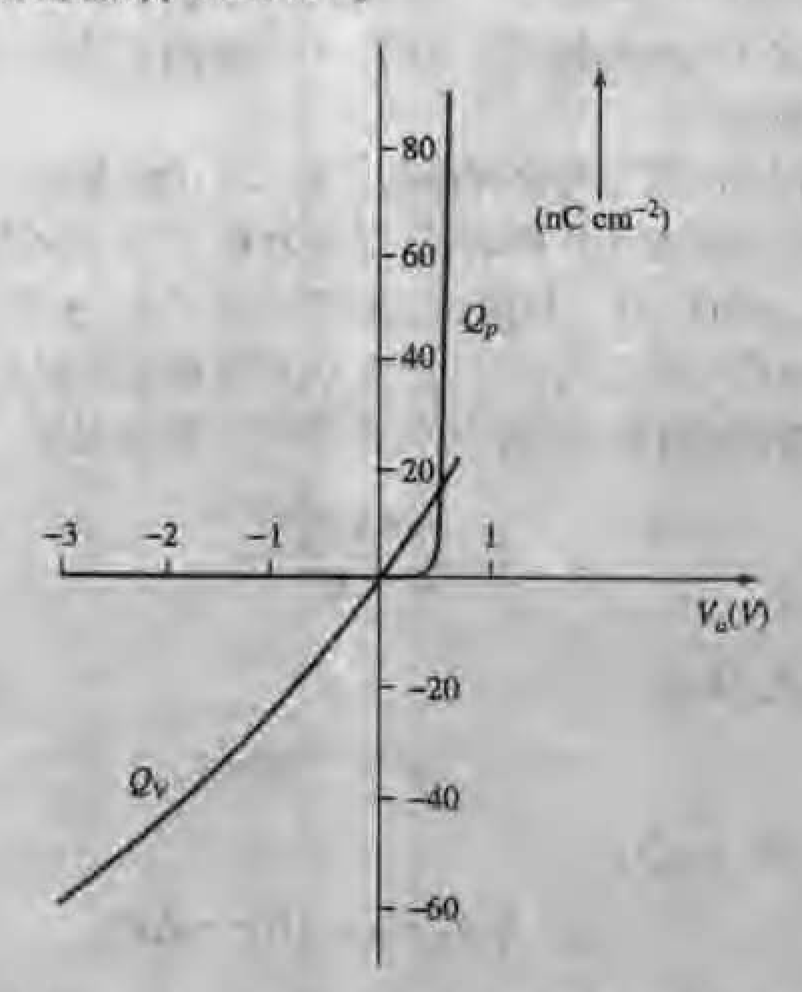
解：对于这种p⁺n二极管，可以忽略注入到重掺杂p型区的电子，利用等式(5.4.8)计算出 $Q_{po} = 6.7\times10^{-16}$C cm^{-2}。与二极管的指数因子相乘得到 Q_p 为

$$Q_p = Q_{po}\left(\exp\frac{qV_a}{kT} - 1\right)$$

Q_V是耗尽区在 $V = V_a$时的存储电荷与 $V = 0$ 时的存储电荷之差的负值。因此,利用等式(4.3.1)得到

$$Q_V = -qN_d(x_d - x_{do}) = -\sqrt{2\epsilon_s qN_d}[(\phi_i - V_a)^{1/2} - \phi_i^{1/2}]$$

这里,由等式(4.2.10)计算出 $\phi_i = 0.872\text{V}$。



利用这两个表达式计算 Q_p 和 Q_V 的值,并绘制在如下同一个图上。我们发现反向偏置时,Q_p和 Q_V相比较可忽略,并且两种存储电荷都是负的。在 0.5 – 0.6V 之间的正向偏置下,两种电荷浓度相等。0.6V 时,Q_p 等于 84.3nC cm^{-2},大约为 Q_V的 3.4 倍。在更高的正向偏压下,Q_p比 Q_V要大很多。这个例题强调了在反偏条件下,耗尽结电荷存储(Q_V)是最重要的,然而在正向偏置时,少数载流子存储(Q_p)占据主导地位。

在结束本节之前,简要讨论一下异质结中的电荷存储。仍考虑宽带隙 n 型半导体和窄带隙 p 型半导体,且假设扩散电流起主导作用。前面看到,正向偏置下,存储在 n 型中性区的过剩空穴 Q_p(单位面积)数量与中性区边界的过剩空穴浓度成正比(等式(5.4.2))。因为中性区边界处的过剩空穴浓度与电流成正比,所以存储的过剩空穴也与电流成正比(等式(5.4.3)和(5.4.5))。

5.3 节讨论的异质结中,注入到 p 型区的电子电流和注入到 n 型区的空穴电流的比值,与同质结有较大的差别,对于给定的电子电流,空穴电流显著减小。因此,异质结中存储在 n 型区的空穴数量(单位面积),比相应存储在同质结的数量少很多。空穴存储的减少使得二极管具有更快的响应速度。异质结的小的电荷存储和快的响应速度,有利于提高异质结双极晶体管(HBT)的瞬态特性,我们将在 7.6 节讨论 HBT 的频率响应特性。

5.5 器件建模和模拟

pn结的电荷存储特性使得用等效电路手工处理包含瞬态问题的计算难度大大增加。例如，集成电路中最常用的二极管开关，在电路操作中不断地正偏和反偏。这种应用中，占主导地位的存储电荷在扩散电荷和耗尽层电荷之间切换。在这种情况下手工计算最有效的方法是，首先确定瞬态过程开始和结束时总的存储电荷，然后，通过用存储电荷的改变 ΔQ 除以二极管开关的驱动电流 I 计算出开关所需的时间（习题5.18）。

这一方法当然不能给出精确的二极管电流和电压关于时间的函数关系，在很多应用中这些细节也不需要。如果需要精确的函数关系，可以用分段线性来近似。“分段线性”这个词意味着非线性电荷存储效应在小的电压范围内可以一级近似成线性元件。如果电压增量取值足够小，这一近似是精确的。对于给定的瞬态问题，只要将足够多的线性近似组合起来，代表整个电压变化过程的分段线性近似便可以达到任意的精度。

5.5.1 集总元件模型

尽管没有明确指出，但是等式(4.3.8)和(5.4.7)中的小信号电容 C_j 和 C_d 是分段线性近似的，因为它们用线性电路元件（电容）表示了总的非线性电荷存储效应。完整的pn结分段线性模型必须包括电导和电容，以表示真实的流过二极管的电流。理想二极管中的电导（单位面积）为（根据等式(5.3.15)）

$$g_d = \frac{\mathrm{d}J}{\mathrm{d}V_a} = \frac{q}{kT}J_0(e^{qV_a/kT}) = \frac{q}{kT}(J + J_0) \tag{5.5.1}$$

二极管总的分段线性电路由 g_d 与 C_j 和 C_d 并联组成，如图5.14所示。因为我们只讨论最基本的pn结，所以图5.14所示的电路没有包括IC平面工艺在制造过程中引入的其他元件成分我们将在5.5节中讨论这个问题。

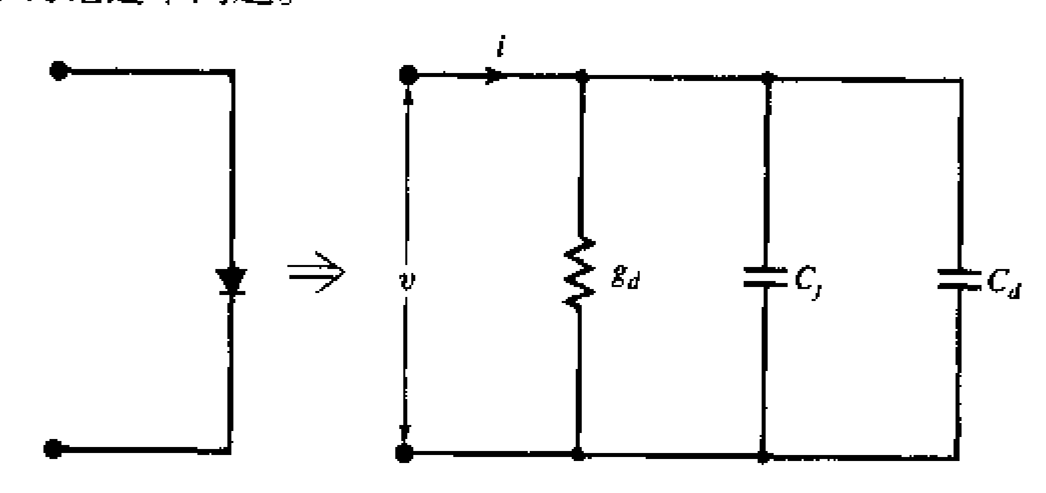


图5.14 二极管小信号等效电路——结型二极管的非线性电流电压关系的分段线性表示。理想长二极管的元件值（单位面积），C_j 由等式(4.3.8)给出，C_d 由等式(5.4.7)给出，g_d 由等式(5.5.1)给出。这一等效电路仅在计算 C_j、C_d 和 g_d 的工作点附近小的偏移范围内有效

对于一些特殊应用，例如在直流偏置的二极管上加小电压正弦激励，利用图5.14给出的电路在某一个工作点上进行分析，能够得到足够的精度。因此该电路被称做二极管小信号等

效电路 二极管小信号等效电路在分析通信电路和线性放大电路等很多电路时都非常有用

利用分段线性的小信号等效电路手工计算分析大信号问题时十分麻烦。除非给大学生作练习,这个问题所有的过程都可以交给计算机解决 为了能在计算机上进行分析,必须给出电流、电压和电容的初值,然后改变一个小量,计算出新的值。

为了解释这种逼近,考虑图5.15所示的电路。初始状态对应于最初的未加偏置的电路(图5.15中 $V_S=0$)。在 $t=0$ 时刻,电压源 V_S 改变. 因为电容上的电压不能跳变,从最初的状态($t=0^-$)可以计算出电容的大小以及代表二极管的电流源的值。假设电路中的元件在小的时间增量 Δt 内保持不变,在 Δt 时刻,二极管上的电压变成新的值 V_D. 这一新的电压值 V_D 被用于计算等效电路中元件的新值。这一过程被重复迭代,直到二极管达到它的终态。

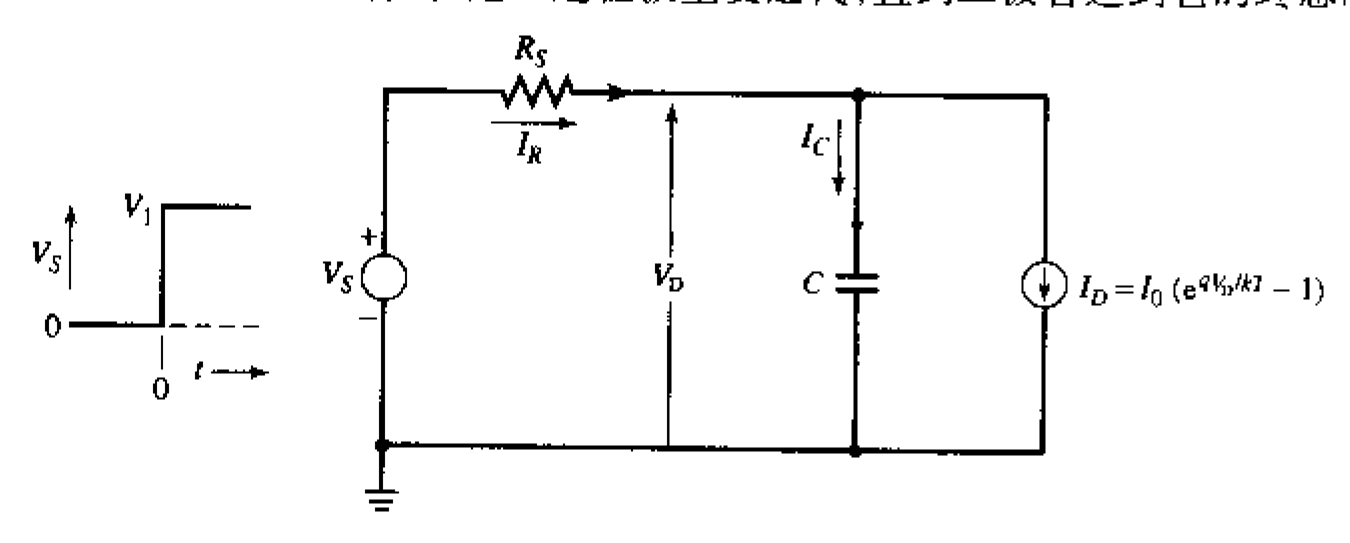


图5.15 典型的分析二极管开关问题的电路。电压源 V_S 在 $t=0$ 时刻从0跳变到了 V_1 关于这个问题的分析见正文

如果图5.15所示的电路通过采用分段线性的计算机程序来分析,整个操作流程如下:

计算机操作	参数值
初始化稳定状态 $V_S=0$ 时的 I_D, C 和 V_D	$V_D=0, I_D=0, C=C_0$
设定时间 $t=0$	$t=0$
设定 V_S 为一个新值	$V_S=V_1$
→ 时间增量	$t=t+\Delta t$
计算电流 I_R 和 I_C	$I_R=(V_S-V_D)/R_S$
	$I_C=I_R-I_D$
计算电压 V_D 在 Δt 期间的改变	$\Delta V_D=I_C\Delta t/C=(I_R-I_D)\Delta t/C$
计算新的电压	$V_D=V_D+\Delta V_D=V_D+(I_R-I_D)\Delta t/C$
计算新的二极管电流	$I_D=I_0(e^{qV_D/kT}-1)$
测试 I_D 是否已经达到稳态值	
没有,继续	
是,从循环退出	
计算新的电容(假设是突变结)	$C=\dfrac{C_0}{\sqrt{1-V_D/\phi_i}}+\dfrac{q}{kT}\tau I_D$
回到循环的最初点,开始下一次时间增量的计算	
→ 停止	

将这一分析过程继续下去，直到计算出的二极管电流与稳定值之差达到设定的精度。计算机指令还包括一些输出命令，可为用户绘制出任意变量，如电压或者电流等与时间的函数关系。

这种分段线性数值逼近的方法可以用在任意非线性电路中，而计算机模拟是研究整个非线性电路性能的惟一合理的方法。分段线性是用计算机求解非线性问题的几种计算方法（或算法）之一。基于分段线性近似，或其他方法的二极管以及其他非线性电路元件的模型通常以子程序的形式包含在更为精确的集成电路计算机分析程序中。在设计和分析集成电路时使用这些计算机模拟程序是非常必要的。

5.5.2 分布式模拟†

上面所讨论的分段连续分析方法，引入了迭代法计算器件性能的思想，并且利用集总元件帮助理解这个概念，但是用迭代方法进行器件模拟以研究器件中电学行为的复杂物理机制，可以提供对器件工作原理更为详细的了解。

我们以常用的器件模拟器 PISCES[10] 为例来讨论器件模拟。PISCES 不仅有广泛用于大学教学的免费软件版本，也有满足工业界要求的商用版本，并有相应的软件包支持。其他模拟器的很多内核与 PISCES 相同。稳态或瞬态条件下，PISCES 通过求解二维空间非均匀网格上的 Poisson 方程（等式(4.1.10)）和一种或两种载流子的电流连续性方程（等式(5.1.3)），来模拟半导体器件的电学特性。通过求解这些方程并采用少量近似，模拟器在工艺和器件开发的初级阶段，对设计工艺实验、理解器件原理和发现潜在问题非常有用[11]。

Poisson 方程根据器件中固定和可动电荷，描述电子的势能分布，可以写成

$$\nabla^2\phi = -\frac{\rho}{\epsilon} = -\frac{q}{\epsilon}(p - n + N_d^+ - N_a^-) - \frac{\rho_F}{\epsilon} \tag{5.5.2}$$

式中 N_d^+ 和 N_a^- 是离化的杂质浓度（$N_d^+ \leqslant N_d$ 和 $N_a^- \leqslant N_a$），ρ_F 是固定电荷浓度（考虑绝缘体时尤其有用）。

电子和空穴的连续性方程描述了载流子浓度

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla J_n - U_n \tag{5.5.3a}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\nabla J_p - U_p \tag{5.5.3b}$$

式中 U_n 和 U_p 是电子和空穴的净复合率，包括了 Shockley-Hall-Read 复合（等式(5.2.9)）和 Auger 复合（等式(5.2.16)）。复合率表达式中的寿命 τ_n 和 τ_p 是载流子浓度的函数。

与 2.8 节讨论的工艺模拟器一样，我们先要进行网格划分。通常网格是非均匀的（一般采用三角形网格），网格上已定义了杂质分布，这个杂质分布是从解析函数给出或者由工艺模拟器如 SUPREM（2.8 节）给出。每一个网格点都与三个方程相联系：Poisson 方程，两种载流子的连续性方程。在求解 Poisson 方程和两个电流连续性方程时，模拟器需要在整个可能包含不同电子材料的区域，自恰求解关于电势 ϕ 和电子空穴浓度 n 和 p 的三个偏微分方程。

考虑 N 个节点，在每一个节点上，首先将连续的解析微分方程转化成差分方程，实现空间的离散化。每一节点都代表包含它的小体积。考虑进入小体积的净电荷流量，及内部的产生

和复合,保持电荷和电流守恒。每个节点的特性主要取决于最近邻的节点。计算出每一点的电势和载流子浓度,然后沿每个三角形线段计算出电流,连接线段的两个端点间的电势差除以线段的长度得到沿每一条线段的电势梯度(电场)。

然而,这3N个代数方程(一个Poisson方程,两个连续性方程)相互耦合,而且是非线性的,所以不能直接求解,需要给定初始猜解通过非线性的迭代方法求解。每种数值方法都需要求解大量的方程系统,总的方程数量近似等于网格数量的1~3倍。

在某些情况下,可以简化分析,缩短计算时间。例如,当没有电流流动时,仅求解Poisson方程即可(N个方程)。在某些情况下,只需要了解一种载流子的行为,这样就可以只求解Poisson方程和一个连续性方程(2N个方程)。这种类型的模拟对分析多数载流子器件是非常有用的,比如MOS晶体管(MOSFET)(第9章和第10章)、JFET(4.5节)和MSFET(4.5节),当然也适用于没有正向偏置结的器件(例如电耦合器件和电容)。然而,即使在像MOSFET这样的多数载流子器件中,少数载流子效应也变得越来越重要。例如,衬底电流和注入到栅氧化层(第10章)的热载流子会严重影响到MOSFET的设计。所以,要仔细分析MOSFET,两种载流子类型都需要考虑(3N个方程)。而双极晶体管的分析(第6章和第7章)必须包括两种载流子类型。

迭代方法的收敛速度严重影响到迭代次数。某些算法线性收敛,每一次迭代的误差以相同的因子减小,某些算法可能收敛更迅速。两种常用的数值求解方法是Gummel法和Newton法。

Gummel法对方程依次顺序求解,先假定载流子浓度固定来求解Poisson方程,将求出的电势代入可直接求解的线性的连续性方程中,得到的载流子浓度被重新代回Poisson方程,作为下一步迭代的初始值。因为每次只求解一个方程,而待求解的矩阵包含N行N列,所以当方程之间的耦合很小时,这种方法很有效。方程之间最大的耦合项是漂移电流项,所以当电流较小时(例如在器件的隔离结构中),Gummel法收敛速度较快;当电流较大时(例如在电阻中),收敛速度较慢。

Newton法在迭代过程中允许改变所有的变量,相互耦合的代数方程组通过矩阵求逆同时求解。Newton法非常稳定,收敛速度几乎与偏置条件无关。然而,因为同时考虑变量电势、电子和空穴,需要求逆的矩阵的行和列分别是单个变量的三倍。在网格相同的情况下,$3N\times3N$矩阵求逆所需的时间是$N\times N$矩阵的20倍左右。如果只考虑Poisson方程和一种载流子类型,那么只需要求解一个$2N\times2N$的矩阵。Newton法一般更适用于求解一种载流子的问题,对于两种载流子问题,Newton法也是可行的,但是如果器件结构复杂,计算所需的时间和内存非常大。另一方面,随着电流的增大,Gummel法收敛速度减慢,甚至在大电流时不收敛。

尽管需要求逆的矩阵很大,但是包含有很多零元素。如果每个结点只与它最近邻的结点有关。矩阵中每个结点只与邻近的四个结点(矩形网格)耦合。这样对Gummel法来说,矩阵的元素为5N。对Newton法来说,如果只考虑一种载流子,矩阵元素增加了4(2×2)倍,如果考虑两种载流子,则增加9(3×3)倍。如果考虑更多的物理效应,或者采用非矩形网格会增加矩阵元素的数量。然而,矩阵的复杂性与结点数成线性关系,而不是平方关系,因为大量的非对角线上的元素都是零。

在模拟过程中,可以对连续性方程进行时间的离散化,求解得到与时间相关的解。考虑时间间隔Δt_k内的变化,并与前一个时间间隔对应的值联系起来,连续性方程(等式(5.5.3))可

近似为

$$\frac{n_{k+1} - n_k}{\Delta t_k} = f_n(\phi_k, n_k, p_k) \tag{5.5.4}$$

$$\frac{p_{k+1} - p_k}{\Delta t_k} = f_p(\phi_k, n_k, p_k) \tag{5.5.5}$$

式中f_n和f_p分别是等式(5.5.3a)和(5.5.3b)所描述的函数。也可以采用更复杂的算法，增量与前几个时间间隔的值有关。

当参数在时间和空间上剧烈变化时，收敛可能会有困难，并且当外加偏置变化很大时，例如，对电极直接施加阶跃电压时，计算中就会出现问题。为了加快收敛速度，计算的初始值需要正确地给出。例如，用前一步的解，或者前一步解的外推值，或者零偏置下的解作为初值。

如果考虑不同材料组成的区域，必须给出界面处的边界条件。可以采用的边界条件有：欧姆接触、Schottky 接触、绝缘边界条件和反射边界条件。欧姆接触的表面势、电子和空穴浓度是确定的，如果电极上有外加电压，则该处的准 Fermi 能级也确定下来，且欧姆接触处不存在空间电荷。Schottky 接触由金属电极和半导体之间的功函数差以及可选择的表面复合速度来定义。绝缘边界条件由决定表面势的功函数来描述。反射边界条件用于被模拟区域的非接触边界，这样电流只能从器件的接触区流进或流出。反射边界条件下，电位移的垂直分量之差，必须等于沿界面的面电荷密度。

当器件变小时，必须考虑二阶物理效应，我们将在本书后面的章节中讨论。这时，模拟必须考虑这些物理效应的模型，才能精确描述器件性能。此外，当集成电路上的器件数量增加时，器件之间的互连变得更加重要。例如，当互连线之间的距离足够小时，它们之间的电容将占整个电容的大部分，不能再被忽略。在某些电路中，互连线的电阻和电容与有源器件一样重要。所以，除了分析有源器件外，模拟工具还应该考虑集成电路互连问题。

在 2.8 节中看到，在器件模拟中，正确划分节点（网格点）很重要，因为它影响到模拟时间和所需要的内存。计算量的增加比网格点的增加快很多，通常以 1.5～2 的幂指数关系增加。有效的网格划分需要在特性改变较快的区域增加节点，减少变化缓慢区域的节点。同样，对于小器件和非平面器件结构，网格的划分必须适应器件的形状，如图 2.40 所示。不规则的网格结构可用于分析任意形状的器件，并且允许网格只在某些区域加密，而不增加其他区域的网格点数。

然而，不规则网格比矩形网格难定义，网格的自动加密技术（"重新划分网格"）对于关键变量的剧烈变化区域非常有用。初始划分的网格可以是粗糙的矩形网格，然后进行网格的自动加密，最后可能转化为三角形网格，直到能够精确地表示结构。当器件的掺杂分布比较复杂时，这一重新划分网格的过程尤为重要。最初，网格可以在预测的掺杂浓度和电势变化剧烈的区域进行加密。在模拟过程中，电荷密度和电势在空间变化加剧，可能会降低数值计算的精度。先进的模拟器可以自动重新划分网格来提高模拟精度。例如，初始的粗糙网格可能先在掺杂突变的区域加密，在其后的模拟中，可根据电势的变化来加密网格。

随着计算机功能的增强，工艺、器件和电路模拟的结合被越来越广泛地用于设计先进的器件和电路，还可减小昂贵的实验费用。器件尺寸的减小和制造工艺的复杂性使得从实验上了解器件的物理特性和工作原理变得很困难。

因为模拟中的很多步骤，比如节点的细化、图形的输出，在工艺、器件和互连的模拟中都要

用到,将这几个模拟器结合可更有效地利用计算机资源,并有助于全面理解器件的性能。工艺模拟器能提供器件中的几何结构和掺杂分布,这些信息作为器件模拟器的输入,由器件模拟器又可得到器件内部状态的信息,以及各电极的电学特性。器件内部的电势、电场和载流子浓度分布能够帮助我们理解器件的性能,改进器件的结构(例如,减小峰值电场)。从各电极的输出特性和器件版图提取的参数可用于电路模拟,获得电路的开关特性和其他的电路性能参数。

5.6 器件

5.6.1 集成电路二极管

在 n 型硅片上扩散一个 p 型区,然后在硅片的正面和反面引出接触电极,如图 5.16 所示,这是利用集成电路工艺制造出的最简单的 pn 结二极管。如果二极管是集成电路的一部分,由于许多 pn 结都被做在同一芯片上,这种简单的制作方法将导致所有的 n 型区相连,显然会影响电路的正常工作。绝缘氧化层与始终处于反偏的 pn 结相结合,可以避免集成电路中二极管之间或者器件之间不必要的直流互连。氧化层通常用于二极管间的隔离,反偏 pn 结用于二极管下方与硅衬底的隔离。这种氧化层和 pn 结相结合的隔离方法应用广泛。下面考虑一个二极管阵列,来说明集成电路中的二极管结构。

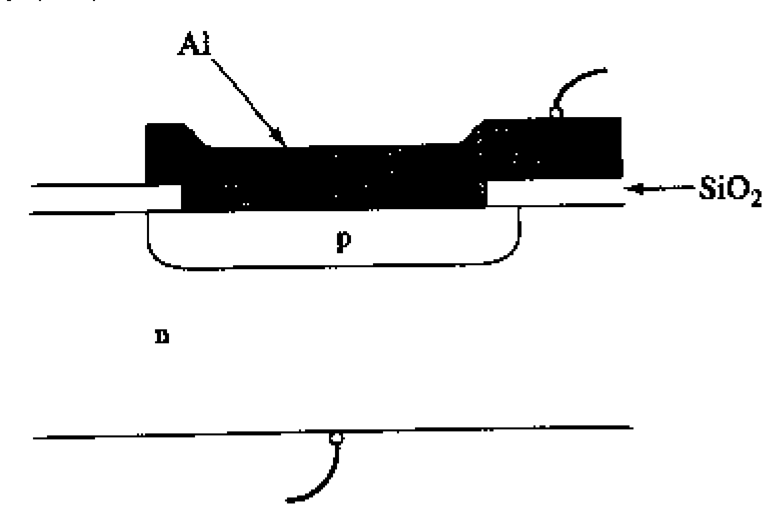


图 5.16 扩散 pn 结在 n 型硅片上扩散 p 型区,表面用二氧化硅钝化,硅片正面和背面都是欧姆接触

为了得到相互隔离的二极管阵列,通常在 p 型衬底上生长一层薄的 n 型外延层,如图 5.17(a)所示。对于数字逻辑开关电路来说,外延层的典型厚度一般是 0.5 ~ 5μm,线性电路约为 5 ~ 20μm(因为工作电压较高)。反偏的外延层-衬底 pn 结提供了垂直方向上的电学隔离。通常是轻掺杂的 p 型衬底材料(典型值为 20Ω · cm,$N_a = 7 \times 10^{14} cm^{-3}$)形成了宽的耗尽层,可减小不必要的有源区与衬底之间的寄生结电容,并保证较高的击穿电压。

完成外延层生长之后的工艺步骤包括光刻、刻蚀和氧化以形成氧化物隔离结构。利用2.6 节介绍的局部硅氧化(LOCOS)工艺来形成氧化物,氧化之前,可以在隔离区对硅外延层进行部分刻蚀。因为氧化物必须穿透整个外延层,如图 5.17(b)所示,LOCOS 仅能提供外延层厚度小于 1μm 的隔离。对于更厚的外延层,可以采用"沟槽隔离"工艺(见 2.6 节)。在隔离区,沟槽刻蚀深入到整个外延层厚度(需要刻蚀进入衬底以适应工艺的变化)。先在沟槽上生长薄的氧化层,然后在沟槽的剩余部分通过化学气相淀积填充氧化物或者多晶硅。沟槽外多余

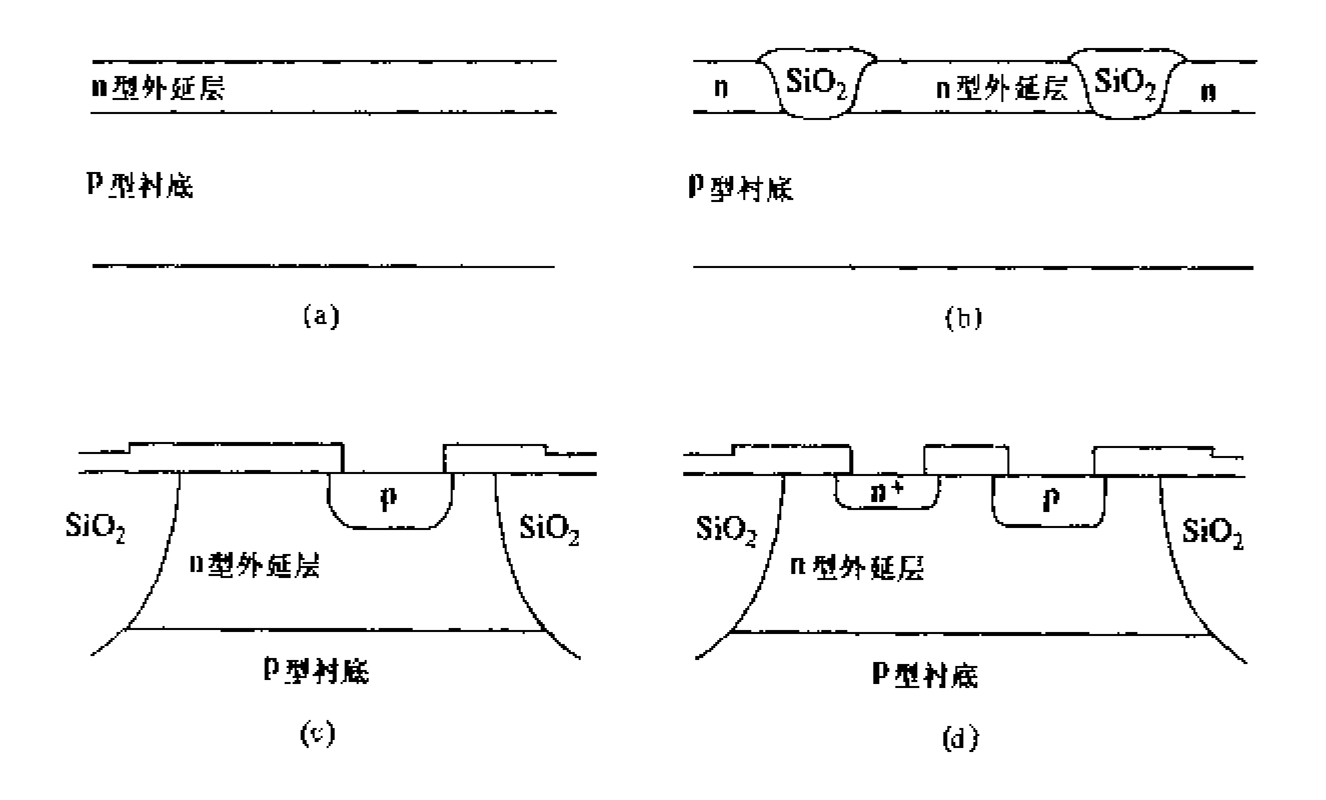


图5.17 集成电路中隔离的结型二极管的平面工艺流程:(a)在p型衬底上生长n型外延层形成衬底结;(b)局部硅氧化(LOCOS)贯穿整个n型外延层,形成横向隔离;(c)在外延层上进行浅p型二极管扩散;(d)在n型外延层上扩散形成高掺杂n型区,提供n型区的欧姆接触(c和d在垂直方向上放大)

的淀积材料通过抛光去除。这种氧化物隔离和pn结隔离结合的技术将形成一个"阱"或者"槽",集成电路器件就制作在其中。

隔离之后的下一步骤是扩散p型区制造pn结二极管。扩散时间不能太长,保证形成的pn结位于n型阱内,如图5.17(c)所示。必须保证二极管的空间电荷区和衬底结的空间电荷区不能接触,这一点很重要。即使在电路工作的"最坏的情况",也必须保证空间电荷区不互相接触。因为反偏电压升高将使空间电荷区变宽,所以应该根据最高的反偏电压来选择外延层厚度和掺杂浓度。两个空间电荷区相遇的情况被称为穿通,在这种条件下,二极管的p型区和p型衬底将不再互相隔离,它们之间将有大的电流流过(习题5.20)。

下一步是在芯片的两端引出二极管的电极。首先,在n型区扩散形成高掺杂的n^+区,如图5.17(d)所示,以获得良好的欧姆接触(由第3章知,金属铝不能和轻掺杂的n型硅形成欧姆接触)。然后光刻接触孔,接着在硅片上覆盖铝,通过选择性刻蚀确定铝的图形,最后进行合金以形成与硅衬底的良好接触,这样二极管就制备完毕(图5.17的垂直方向的尺度被放大了)。

然而,上述流程制造的集成电路二极管有一个严重的不足。为了有较大的传输电流面积,希望二极管电流在垂直方向流动。这样,作为多数载流子的电子需要穿过两个pn结之间的一个又长又窄的通道。这个通道非常类似于第2章末介绍的夹断电阻。对于典型的集成电路来说,外延层电阻率一般是0.5Ω·cm,那么这条通道的串联电阻大约是几千欧。即使是在相对较低的正向电流下,这么大的串连电阻也会造成相当大的欧姆压降,引起pn结的偏置沿这个方向改变。因为电流与正偏结电压呈指数关系,在接近n^+电极的结区会出现电流集边现象,如图5.18(a)所示。这样,二极管的有效面积将小于它的几何面积,导致很多不必要的效应产生。

为了消除这个问题,可在生长外延层之前,在衬底上制作重掺杂的n型区,如图5.19(a)所示,这个区域通常处于结区的下方。在外延层生长、氧化、扩散等步骤完成之后,得到如图

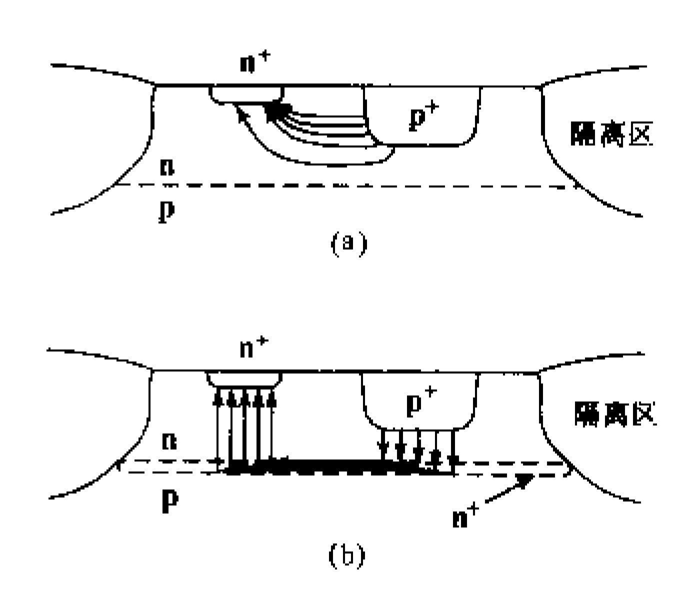


图 5.18　(a) 大的外延层串联电阻将导致电流在 pn 结接近 n^+ 欧姆接触区域出现"集边"现象;(b) 低电阻埋层使得电流在整个二极管区域流过均匀

5.19(b)所示的结构。采用这一 n 型的埋层,串联电阻降低到只有几个欧姆,这通常是可以接受的,电流在整个二极管面积上基本上均匀流过,如图 5.18(b)所示。埋层的另一个好处是减小了表面和 p 型衬底之间穿通的可能性,因为埋层中的施主抑制了空间电荷区向 n 型外延层的延伸。虽然抑制了穿通,但同时降低了衬底结二极管的击穿电压,因为埋层的出现减小了空间电荷区宽度,增加了最大电场。

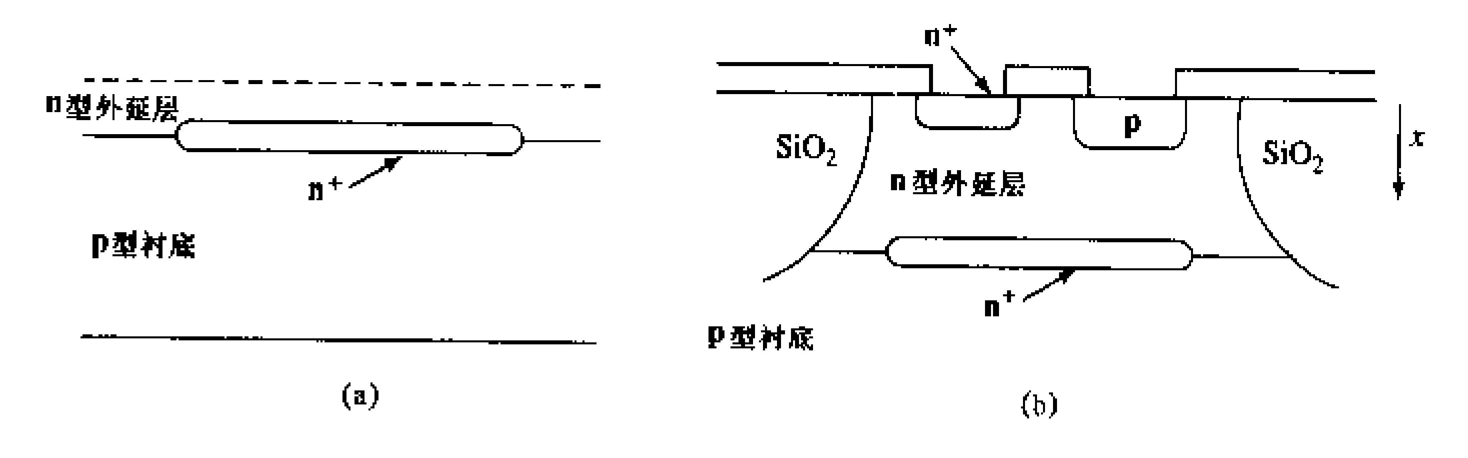


图 5.19 (a)在生长外延层之前,在 p 型衬底上扩散形成重掺杂的埋层,这将减小二极管串联电阻;(b)具有埋层的集成电路二极管

埋层的引入增加了集成电路制备工艺的复杂性。首先,因为埋层的杂质是在外延层生长之前引入的,这些杂质要经历外延生长和其后各种高温工艺,这意味着必须采用扩散缓慢,相对不易挥发的杂质。缓慢扩散是为了保证器件所要求的特征尺度;不易挥发是为了避免生长外延层时对外延层的自掺杂。实际工艺中使用砷或锑作为埋层杂质可以获得较好的结果。然后,采用扩散系数很大的硼形成 p 型区,其后的浅 n 型扩散采用砷。因为缺少类似砷和锑的扩散缓慢的 p 型杂质,所以制造实用的 p 型埋层很困难。

因为 n 型埋层使得 n 型区和 p 型衬底之间的空间电荷区变窄了,所以增加了外延层和衬底之间的寄生电容,这是仅在有源区下方引入埋层的原因之一。引入埋层的一个不可避免的缺点是高掺杂带来的应力使外延层载流子的寿命和迁移率降低。这一区域是决定器件性能的关键区域,所以问题可能很严峻,采用砷比采用锑的效果要好一些。需要综合考虑埋层带来的优缺点以优化设计。

下面我们介绍一种集成电路设计和分析中被广泛采用的表格和列线图。在均匀掺杂的衬

底上进行高斯分布和余误差分布掺杂,形成的两种pn结的性质在Lawrence和Warner的书[12]中已经讨论过了。他们利用计算机程序精确求解了pn结方程,计算得到了结电容,总的耗尽区宽度以及结任意一侧的耗尽区宽度。文献[12]提供了非常有用的数据。要利用Lawrence和Warner的结果,需要知道整个区域内的背景掺杂浓度,以及进行扩散的表面掺杂浓度。这些数据以总的结电压为变量确定了一组输出曲线,而且他们的分析还给出了最大电场与总的结电压之间的关系。参考这些结果对设计集成电路起到了辅助作用,这些结果也可以作为精确数值模拟的初值。

最后我们注意到,虽然本节中提到的LOCOS隔离技术,经常被用于IC设计,但是在小尺寸电路中已经开始采用新的技术。例如,2.6节介绍的在器件之间不采用LOCOS工艺,而是直接制作绝缘硅氧化物的沟槽隔离技术。这种新技术显著减少了隔离区的面积。

5.6.2 发光二极管

在有效的正向偏置二极管(例如,空间电荷区内几乎没有复合中心)中,载流子从二极管一侧的中性区的注入,穿越变窄的空间电荷区进入另一侧中性区,在那里变为少数载流子,它们边扩散,边和多数载流子复合。间接带隙半导体,例如硅中的电子与空穴的复合需要伴随动量的改变和能量的损失,这样电子与空穴的直接复合不容易发生,而是通过带隙中间的复合中心复合(见5.1节)。

然而,在直接带隙半导体中,比如砷化镓,当电子从导带跃迁填充价带中的空位时,电子和空穴直接复合,动量不发生变化。适当地设计半导体结构,这一跃迁所释放的能量通常以光子的形式发射出去。这种类型的发光二极管(LED)应用广泛。红色的LED通常用于指示器和普通家用电器上的小型显示器上(例如钟表)。发射的光的颜色与半导体的带隙有关,最常用的直接带隙半导体发红光,或者红外光($E_g \approx 1.5 \sim 1.8\text{eV}$)。为了得到能量更高的蓝光和绿光,宽带隙的基于氮化镓以及与其相关的化合物已经研制出来了。如果LED的光谱能覆盖整个可见光范围,由于其节能的特点,LED的应用范围将更广,例如用在交通灯上,不仅能节约电能,而且由于其寿命很长(几年),降低了成本。

例题 LED的效率

已知LED每秒钟能产生2×10^{15}个光子,发光波长为0.8μm,如果外加1V电压将产生1mA电流,计算LED的效率(光能输出/电能输入)。

解:每个光子的能量是$1.24\text{eV}/\lambda(\mu\text{m}) = 1.24\text{eV}/0.8 = 1.55\text{eV} = 2.48\times10^{-19}$焦耳。总的发射出的光所对应的能量是$2.48\times10^{-19}\times2\times10^{15}$焦耳/秒$=5\times10^{-4}$瓦。提供给LED的电能是$10^{-3}$瓦,所以效率就是$0.5=50\%$。

小结

可用于求解自由载流子浓度变化的偏微分方程,即连续性方程,是分析半导体器件的有力工具。在pn结的准中性区附近求解少数载流子的连续性方程,可以得到载流子的浓度分布。准中性区边界处的少数载流子浓度,可以写成外加结偏压的函数。利用这些边界值,以及电流

与自由载流子浓度的关系,可以得到稳态时的电流电压方程,这就是理想二极管方程 $J = J_0 [\exp(qV_a/kT) - 1]$

通过求解与时间相关的连续性方程可以理解过剩载流子寿命的含义。在大部分情况下,决定载流子寿命的复合率与复合中心性质有关。这些复合中心由局域的电子能态组成,能量位于禁带中,通常接近本征 Fermi 能级。通过局域化中心复合的理论被称做 Shockley-Hall-Read(SHR)复合理论。在反偏和低正偏的硅二极管中,通过空间电荷区复合中心的产生与复合对于电流有很大贡献。Auger 复合发生在载流子浓度很高的情况下。在 Auger 复合中,复合释放出来的能量和动量,将传递给第三个自由载流子。硅表面的局域态和空间电荷区将导致重要的表面复合效应,表面复合通常用表面复合速度 s 来描述。

二极管的瞬变特性受到存储的少数载流子电荷以及耗尽层存储电荷的影响。这些电荷分量与电压是非线性关系,只有在小信号时,才可用线性电路来近似。小信号等效电路中用电容表示电荷的存储。这个电路可通过分段方式得到大信号瞬态特性的精确解。实际问题通常都采用计算机求解。

pn 结二极管是集成电路的基本结构单元。用氧化物隔离器件,可以避免集成电路器件之间不必要的互连,这一技术已得到广泛的应用。表面电极和正向偏置 pn 结之间的串联电阻可以通过在扩散结下方增加掺杂施主的埋层来减小。很多的集成电路 pn 结性质可以参考已经发表的计算曲线。

参考文献

1. K. Thiessen and G. Zech, *Phys. Stat. Sol.* (a) **10**, K 133 (1972).
2. W. Fahrner and A. Goetzberger, *J. Appl. Phys.* **44**, 725 (1973).
3. W. Shockley and W. T. Read. *Phys. Rev.* **87**, 835 (1952).
4. R. N. Hall, *Phys. Rev.* **87**, 387 (1952).
5. E. M. Conwell., *Proc. IRE* **46**, 1281 (1958).
6. J. L. Moll, *Physics of Semiconductors*, McGraw-Hill, New York, 1964, p. 117.
7. P. U. Calzolari and S. Graffi, *Solid-State Electron.* **15**, 1003 (1972).
8. B. L. Sharma and R. K. Purohit, *Semiconductor Heterojunctions*, Pergamon Press, Oxford, 1974, Sec. 1.1.
9. H. Unlu and A Nussbaum. Solid-State Electron. **30**, 1095 (1987).
10. R. W Dutton and Z. Yu, *Technology CAD: Computer Simulation of IC Processes and Devices*, Kluwer Academic Publishers, Boston, 1993.
11. K. M. Cham, S.-Y. Oh, D. Chin, and J. L. Moll, *Computer-Aided Design and VLSI Device Development*, Kluwer Academic Publishers, Boston, 1986.
12. H Lawrence and R. M. Warner, Jr., *Bell Syst. Tec J.* **34**, 105 (1955).
13. P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. F. Gibbons, *Physical Electronics and Circuit Models of Transistors*, Wiley, New York, 1964, p. 75.

参考书

C. M. Wolfe, N. Holonyak, Jr., and G. E. Stillman, *Physical Properties of Semiconductors*, Prentice-Hall, Englewood Cliffs, NJ, 1989, Chapter 9: Heterostructures.

H. Kroemer, Chapter 4 of *VLSI Electronics Microstructure Science. Volume 10: Surface and Interface Effects in VLSI*, ed. N. G. Einspruch and R. S. Bauer, Academic Press. Orlando, 1985.

G. W. Neudeck, *The PN-Junction Diode*, Volume II Modular Series on Solid-State Devices, Addison-Wesley Inc, Reading, MA, 1983.

习题

5.1*　已知 $0.6\Omega \cdot cm$ 的 n 型硅在本征 Fermi 能级处的复合中心浓度为 $10^{15} cm^{-3}$，假定 $\sigma_n = \sigma_p = 10^{-15} cm^2$，$v_{th} = 10^7 cm\ s^{-1}$

(a) 如果该区域可动载流子全部耗尽，计算产生率。

(b) 当少数载流子浓度略低于平衡态时，计算该区域的产生率。

5.2*　光照射在施主掺杂浓度为 $10^{16} cm^{-3}$ 的硅上，均匀产生出 $10^{21} cm^{-3} s^{-1}$ 的电子-空穴对。E_i 处的复合中心浓度为 $10^{15} cm^{-3}$，电子和空穴的俘获截面为 $10^{-14} cm^2$。

(a) 计算光照下稳态的空穴和电子浓度

(b) 在 $t=0$ 时刻，移去光源　计算总的空穴浓度与时间的关系，并求出寿命。假定热运动速度等于 $10^7 cm\ s^{-1}$，没有电流流动

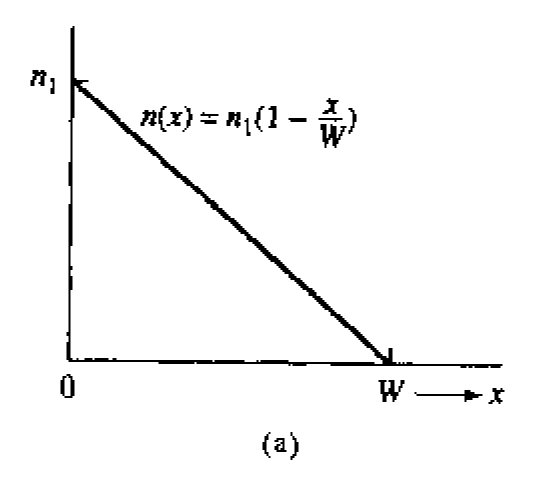
5.3　(这个问题将提供半导体中电中性的概念) 假定没有任何的中和电荷，硅材料中的自由电子可保持如图 P 5.3a 所示的分布　考虑与电荷相关的电场在 $x = W$ 处为零，令 $n_1 = 10^{18} cm^{-3}$，$W = 1\mu m$，$D_n = 7 cm^2 s^{-1}$，$T = 300K$

(a) 计算为维持这一自由电荷分布，$x=0$ 处的电场和电流　(注意漂移分量和扩散分量)

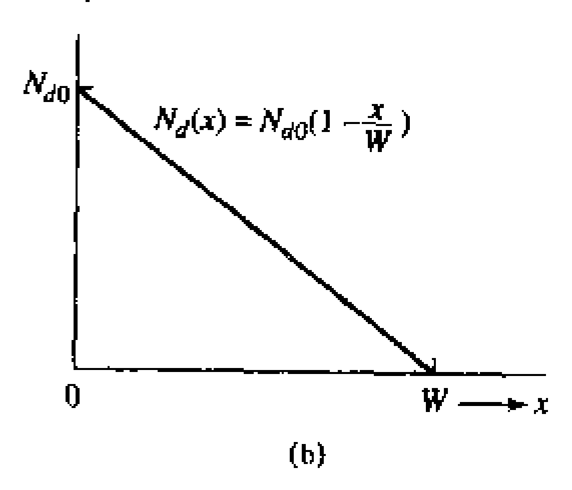
(b) 依据 (a) 的结果，分析这个电荷分布是否合理。

(c) 如果硅中具有均匀的电子浓度 n_1，并维持电流密度近似为 $10^5 A\ cm^{-2}$，假设此时硅中的功耗不至于造成不可逆的破坏，求可能存在的最大电场　(实际的 IC 器件中电流密度比这个值低两个或者更多的数量级)

(d) 如果硅中存在如图 P5.3b 所示的补偿正电荷浓度分布，当电场取 (c) 中的值时，计算 N_{d0} 的值，用自由电子浓度 n_1 来表示，分析硅是否接近电中性 (利用 $x=0$ 处的值)。



(a)



(b)

图 P 5.3

5.4†　如果电子通过类施主复合中心复合，俘获截面 σ_n 可以这样粗略估算：假定某一区域的自由载流子的热能 $3kT/2$ 小于库仑作用能量，这时该区域的载流子可被复合中心吸引过来。这个区域所涉及到的面积就是 σ_n。

(a) 根据这个模型求取 σ_n 的表达式，估算硅在 300K 时的值。

(b) 分析该模型中 σ_n 与温度的关系，以及其他与温度有关的量。复合中心的电子俘获率与温度又是什么关系？(假设非本征硅中存在复合中心)

5.5†　5.2 节在分析通过复合中心的复合时，曾经假定小注入，且没有电流流过，这样就可以得到过剩载流子衰减的特征时间。根据类似的分析，求相反情况，即大注入时 n 型半导体中的"有效寿命"，假设电中性成立。令 $n' = p' \gg n_o$ 和 p_o，当 $\sigma_n \neq \sigma_p$ 和 $\sigma_n = \sigma_p$ 时，与小注入时的衰减时间或有效寿命进行比较。

5.6　在短二极管中考虑穿越 pn 结进入 n 型区的少数载流子空穴的连续性方程，假设空穴的复合基本上发生在电极 $x = W_B$ 处。直接利用等式 (5.3.10) 证明空穴分布是线性的，与等式 (5.3.17) 一样。

5.7 考虑理想的长二极管,忽略空间电荷区的产生与复合,这样电流电压关系就可用等式(5.3.15)来描述。证明等式(5.3.15)描述的反偏电流等于 pn 结任意一侧的少数载流子的产生率的积分。提示:利用与等式(5.3.12)对应的电子的表达式,考虑 $G_p - R_p = -p'/\tau_p$

5.8 空穴通过正向偏置的 pn 结注入到 n 型区,n 型区长度远大于空穴扩散长度 L_p。利用等式(5.3.12)证明,L_p 是空穴在 n 型区复合之前扩散的平均长度。

5.9* 考虑正向偏置的理想 pn 结二极管中的电流,证明空穴电流与总电流的比值可以通过改变结两侧的相对掺杂浓度来控制。如果将空穴电流与总电流的比值称为 γ,将 γ 表示成 N_a/N_d 的函数。已知 n 型区电阻率等于 0.001Ω·cm,p 型区电阻率等于 1Ω·cm,计算 pn 结的 γ。假设 $\tau_p = 0.1\tau_n$,两个中性区长度都远大于各自的少数载流子扩散长度。

5.10† 已知 pn 结二极管结构如图 P 5.10 所示。假设

(1) $N_a = N_d = N_o \gg n_i$;

(2) $W \ll L$, L 是少数载流子的扩散长度, $L_1 \gg L$;

(3)所有的扩散系数等于 D,所有的寿命等于 τ;

(4)空间电荷区宽度远小于 W;

(5)外加电压是 V_B, $V_B \gg \phi_i$, ϕ_i 是内建电势;

(6) $x = +W$ 处的复合率无穷大。

如果一束光照射在二极管 p 型区一侧 $x = -W/2$ 处的平面上,单位时间单位面积产生的空穴-电子对为 G_0。

(a)假设小注入,求二极管中性区的少数载流子浓度分布并绘图。

(b)利用二极管性质和 G_0,计算被光照的二极管中流过的电流。

(c)当光束被移去后,用二极管的相关参数,表示电流的值(稳态)[13]。

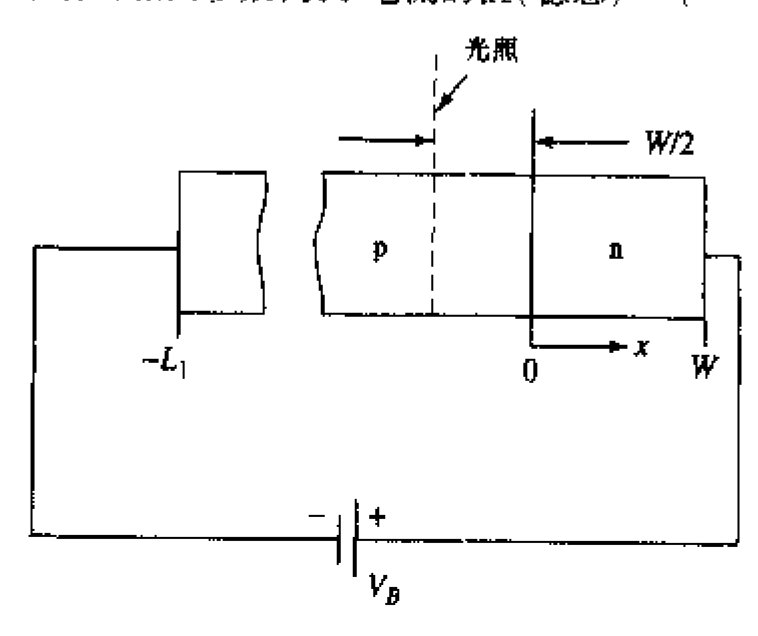


图 P 5.10

5.11* 考虑理想的横截面和掺杂均匀的硅突变长二极管,已知 p 型区和 n 型区的电阻率分别为1Ω·cm 和0.2Ω·cm,少数载流子寿命分别是 $\tau_n = 10^{-6}$s 和 $\tau_p = 10^{-8}$s("理想"表示空间电荷区的效应可以忽略,少数载流子在电荷中性区的流动仅由扩散机制决定)。

(a)求内建电势。

(b)当外加电压为 0.589V($23 \times kT/q$)时,计算空间电荷区边界处的少数载流子浓度。

(c)外加电压与(b)相同,绘出在结两侧多数载流子和少数载流子浓度与位置的关系曲线。

(d)外加电压与(b)相同,求多数载流子电流与少数载流子电流相等的平面位置。

5.12 对于空间电荷区具有较大复合率的 pn 结二极管(正向偏置),绘出中性区空穴电流和电子电流分布。假设

(a)注入的空穴电流是注入的电子电流的两倍。

(b)空间电荷层中电子-空穴对的净复合率是 p 型区中电子净复合率的一半。

证明总的二极管电流等于空间电荷区边界上的总扩散电流与空间电荷区的复合电流之和.

5.13 考虑习题4.10中所示的集成电路的横截面图

(a)忽略串联电阻,但是考虑所有的结,绘出两组 $I-V$ 特性曲线,已知(i)电压 V_C(两种极性)加在电极C上,电极A接地,电极B悬空,(ii)电压 V_B(两种极性)加在电极B上,电极A接地,电极C悬空

(b)* 如果在上述a(i)的条件下,有可能完全耗尽 n^- 层,你认为通过电极A的电流会发生什么情况?为什么?

5.14† 已知均匀掺杂的硅突变pn结二极管面积为 $10^{-5}cm^2$,p型区的受主浓度是 $5\times10^{16}cm^{-3}$,n型区施主浓度是 $10^{15}cm^{-3}$. 采用最简单二极管的理论,即只考虑一维情况,正向偏置下,体复合占主导作用,少数载流子仅通过扩散从空间电荷区边界进入中性区,忽略空间电荷区的复合与产生,每个区的少数载流子寿命是100ns.

这个结型二极管用在如下电路中,要求正向电流 I_F 和反向电流 I_R 的整流比($|I_F/I_R|$)在0.5V时等于 10^4,最大的反向饱和电流为100nA. 二极管能够正常工作的最高温度是多少?仅考虑最重要的温度依赖关系(也就是忽略 D、μ、τ_p、N_c 和 N_v 随温度的改变).

5.15 已知 $\sigma_n=\sigma_p$,证明pn结空间电荷区的复合率在 $p=n$ 且取等式(5.3.23)给出的值时,达到最大

5.16† 利用等式(5.3.25)及其后面一段正文给出的结型硅二极管的值,在半对数坐标系中绘制出正向偏置下 J_r/J_t 与外加电压 V_a 的关系曲线. 已知二极管为n型区高掺杂的突变结(即 $N_a \ll N_d$).

5.17 考虑突变二极管,已知 $(\phi_i-V_a)=5V$,$N_a=10^{17}cm^{-3}$,$\tau_n=10^{-6}s$,$N_d=10^{18}cm^{-3}$,$\tau_p=10^{-8}s$

(a)利用等式(5.3.29)求比值 I_r/I_g(令 $x_i=x_d$).

(b)讨论这个比值与温度的关系.

5.18† 已知理想的突变结基极二极管中,$N_d\gg N_a$,$N_a=10^{17}cm^{-3}$(假设n型区简并掺杂). p型区空间电荷区边界与电极之间的距离是3μm,所有复合发生在电极上. 二极管面积是 $10^{-5}cm^2$.

(a)如果流过二极管的电流是0.5mA,计算p型中性区存储的电荷.

(b)计算正向偏置下空间电荷区存储的电荷

(c)外加0.5mA的电流源时,需要多长时间二极管才能从关断状态($V_a=0$)达到0.5mA的稳态

5.19 考虑处于正向偏置的Schottky二极管的小信号等效电路. 比较这一电路与图5.14所示的pn结二极管的小信号等效电路的异同点.

5.20* 按下述情况,定性分析图5.17(d)所示的结构发生穿通时的效应:(a)沿p型扩散区中部指向衬底的坐标轴方向,画出热平衡时能带图,(b)考虑衬底接地,上方的扩散pn结上加载足够大的反偏电压,使衬底pn结处于正向偏置. 当这一情况下,衬底的空穴会注入到上面的p型扩散区中. 画出这种情况下的能带图.

5.21* 已知理想的硅长二极管的参数为,$N_d=10^{18}cm^{-3}$,$N_a=10^{16}cm^{-3}$,$\tau_n=\tau_p=10^{-8}s$,$A=10^{-4}cm^2$,$T=300K$,计算以下情况的小信号电阻和电容:

(a)正向偏压为0.1V,0.5V和0.7V;

(b)反偏电压为0,5V和20V;

(c)如果p型准中性区的长度是0.1cm,求该区的串联电阻. 分析实际二极管的瞬态响应时,理想二极管模型中必须加入这个串联电阻.

5.22 pn结最初处于关断状态,施加阶跃电流使其开启. 从物理上解释为什么空间电荷区宽度的改变比准中性区少数载流子分布趋于平衡需要的时间更短.(也就是说,比较这两种情况的物理过程)

5.23 如图P5.23所示的硅集成电路二极管中,考虑从p型扩散区进入到 n^+ 埋层的空穴流,用长二极管近似还是用短二极管近似比较好?为什么?并分析从p型扩散区边界到 n^+ 电极的侧向电流. 已知n型外延层的空穴寿命是1μs,n^+ 埋层是过剩载流子有效的复合区.

5.24 高电流密度下,二极管上的外加电压可能主要降落在器件的中性区. 考虑单边突变结的电流与电压关系,轻掺杂一侧的施主浓度是 N_d. 对于典型的集成电路二极管,当10%的外加电压 V_a 降落在n型中性

区上时求电流和 V_a。已知二极管横截面积是 $10^{-5}\mathrm{cm}^2$，n 型中性区的长度是 10μm，$N_d = 5\times10^{15}\mathrm{cm}^{-3}$，$\tau_p$ = 1ns

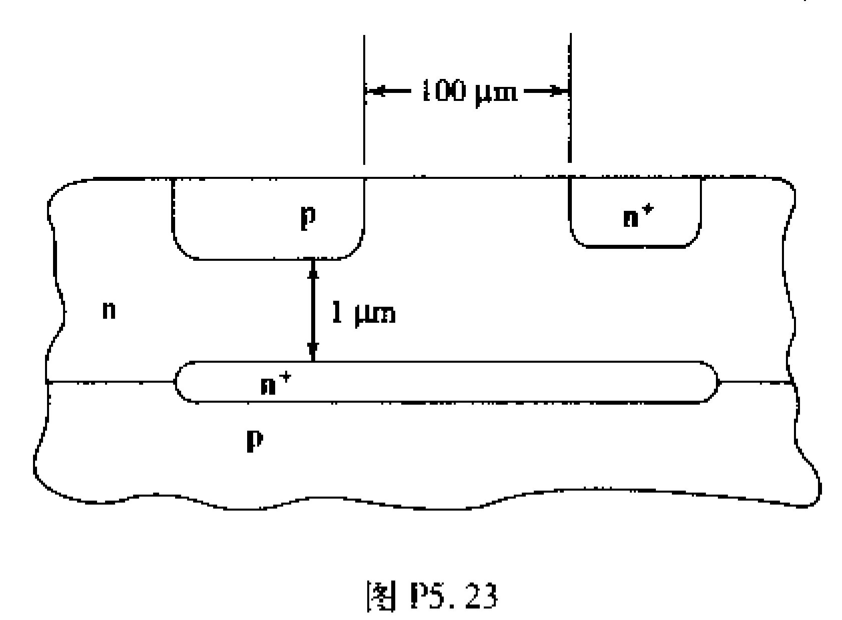


图 P5.23

第6章　双极晶体管I:基本特性

从第5章我们了解到,pn结的p型区相对于n型区正偏时,空穴从p型区注入到n型区,电子从n型区注入到p型区。由于是多子提供电流,随着正偏电压的升高和结势垒高度的下降,电流将迅速上升。反向偏置时,电流由空间电荷区内或附近区域的少子产生流形成,其数值要小得多。不过,如果结附近提供的少子增加,流过反偏结的电流也会上升,例如,二极管光电探测器或传感器中高能粒子的辐射可以导致结附近的少子增加。

另一个增加反偏pn结附近少子数量的方法是在其附近放置一个正偏pn结,该方法的优点是少数载流子数量可以由正偏pn结的偏压控制。

通过改变一个pn结的偏压调制另一个pn结电流的方法称为双极晶体管效应。这是器件电子学历史上最重要的思想之一,William Shockley、John Bardeen和Walter Brattain由于发明了双极晶体管(BJT),获得了诺贝尔物理奖。

这一章中我们将分析双极晶体管的基本工作原理。为了简化分析,我们将分别考虑晶体管各种特定的工作区。首先考虑晶体管有放大作用的放大区,然后在介绍晶体管开关特性时讨论其他工作区。随后介绍Ebers-Moll模型,该模型能够描述晶体管在各个偏置区的基本工作原理,然后将这个模型与本章开头关于晶体管原理的讨论联系起来。接下来,我们将考虑平面扩散双极晶体管(放大晶体管与开关晶体管)的设计。最后讨论不同区域半导体材料不同的异质结双极晶体管(HBT)。

6.1　晶体管工作原理

我们用图6.1所示的简单结构分析晶体管的工作原理[1]:在一块半导体材料上有两个间距为W,横截面积为A的pn结。两个结靠得"很近",使得在V_{BE}为正时通过结1注入的电子能够到达结J_2附近,也就是说W足够小,几乎没有电子在中间的p型区被复合。这个中间区域称为晶体管的基区。

注入结(或发射结)附近的n型区被称做发射区,集电结附近的n型区被称做集电区。我们主要研究电子如何从发射区注入到基区,并流经基区到达集电结。我们先不考虑基区电子的复合或产生作用,该作用的影响将在后面讨论。

图6.1所示的双极晶体管中,还存在从结J_1流向结J_2,或者从结J_2流向结J_1的空穴(基区多子)电流,这部分电流可以忽略,因为在任何偏置条件下,从任一n型区流向p型基区的空穴流都非常小。如果把垂直结平面的方向设为x,可以从等式(1.2.22)导出沿x方向的空穴电流密度表达式(不考虑复合)

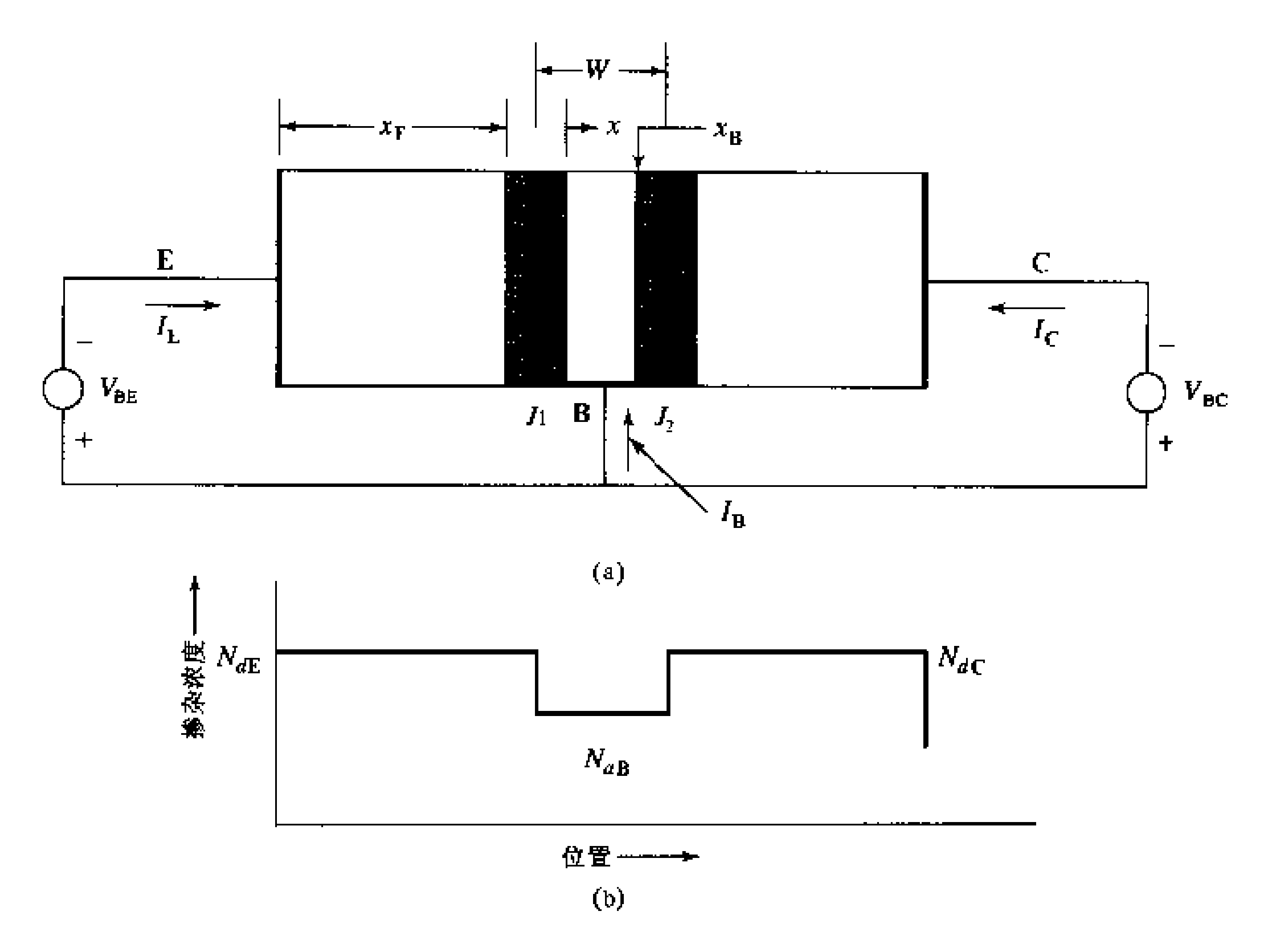


图 6.1 (a)原型晶体管,pn 结 J_1 和 J_2 的间距为 W,(b)各区均匀掺杂,掺杂浓度在结 J_1 和结 J_2 处突变,中间 p 区的准中性部分被空间电荷区分别限制在 $x=0$ 和 $x=x_B$ 之间.电极 E、B、C 为欧姆接触

$$J_p = 0 = q\mu_p p\mathscr{E}_x - qD_p\frac{\mathrm{d}P}{\mathrm{d}x} \tag{6.1.1}$$

和

$$\begin{aligned}\mathscr{E}_x &= \frac{D_p}{\mu_p}\frac{1}{p}\frac{\mathrm{d}P}{\mathrm{d}x} \\ &= \frac{kT}{q}\frac{1}{p}\frac{\mathrm{d}P}{\mathrm{d}x}\end{aligned} \tag{6.1.2}$$

由于 x 方向的基区空穴电流等于零,可以推导出 x 方向的电场的表达式(6.1.2)①。这个电场不仅与多数载流子(空穴)浓度有关,还与其梯度有关。与空穴电流相比,两个结之间存在电子电流的可能性更高,因为只要结正偏,两边的结都能够稳定地从 n 型区提供电子给 p 型区。根据等式(1.2.21),得到电子电流密度

$$J_n = q\mu_n n\mathscr{E} + qD_n\frac{\mathrm{d}n}{\mathrm{d}x} = kT\mu_n\frac{n}{p}\frac{\mathrm{d}P}{\mathrm{d}x} + qD_n\frac{\mathrm{d}n}{\mathrm{d}x} \tag{6.1.3}$$

将 Einstein 关系 $D_n=(kT/q)\mu_n$ 代入等式(6.1.3),得到

$$J_n = \frac{qD_n}{p}\left(n\frac{\mathrm{d}P}{\mathrm{d}x} + p\frac{\mathrm{d}n}{\mathrm{d}x}\right) = \frac{qD_n}{p}\frac{\mathrm{d}(np)}{\mathrm{d}x} \tag{6.1.4}$$

① 严格地说,等式(6.1.1)与(6.1.2)为近似结果,稍后要讨论的二阶效应在 x 方向能够产生很小的空穴电流。

等式(6.1.4)适用于任意掺杂分布。为了更直观地研究晶体管特性,考虑一种简单的晶体管结构,如图6.1所示:pn结是突变结,基区由一种材料组成且掺杂均匀。这种结构称为原型晶体管,与W. Shockley最初分析双极晶体管时所描述的器件类似[1]。

基区均匀掺杂时多数载流子浓度 p 接近于常数。与 p 相比,少数载流子浓度 n 非常小,这样等式(6.1.4)可以简化为

$$J_n = qD_n\frac{\mathrm{d}n}{\mathrm{d}x} \tag{6.1.5}$$

进一步假设基区的电子复合电流很小(类似5.3节介绍的短二极管),这样基区内 n 线性变化,由此可以得到

$$\frac{\mathrm{d}n}{\mathrm{d}x} \simeq \frac{n_p(x_B) - n_p(0)}{x_B} \tag{6.1.6}$$

用掺杂浓度表示少数载流子电子的浓度,得到

$$J_n = \frac{qD_n n_i^2}{x_B N_{aB}}\left[\exp\left(\frac{qV_{BC}}{kT}\right) - \exp\left(\frac{qV_{BE}}{kT}\right)\right] = J_S\left[\exp\left(\frac{qV_{BC}}{kT}\right) - \exp\left(\frac{qV_{BE}}{kT}\right)\right] \tag{6.1.7}$$

式中

$$J_S = \frac{qD_n n_i^2}{x_B N_{aB}} \tag{6.1.8}$$

从等式(6.1.7)可以看出,通过改变结电压,可以控制电流 J_n。如果 V_{BC} 和 V_{BE} 都是负的,并且远大于 kT/q,那么 J_n 非常小。另一方面,如果 V_{BC} 或 V_{BE} 为正并且大于 kT/q,J_n 主要决定于数值较大的那个正电压。

在分析原型晶体管之前,我们注意到在大多数集成电路晶体管中,基区的掺杂浓度随位置变化。等式(6.1.4)右边的式子为全微分,适用于更普遍的情况

$$J_n = \frac{qD_n}{p}\left(n\frac{\mathrm{d}p}{\mathrm{d}x} + p\frac{\mathrm{d}n}{\mathrm{d}x}\right) = \frac{qD_n}{p}\frac{\mathrm{d}(pn)}{\mathrm{d}x} \tag{6.1.9}$$

取任意的积分限 x 和 x',可以写出等式(6.1.9)的积分形式。忽略基区复合作用,则 J_n 可以提到积分号外

$$\begin{aligned} J_n\int_x^{x'} \frac{p}{q}\frac{\mathrm{d}x}{D_n} &= \int_x^{x'} \frac{\mathrm{d}(pn)}{\mathrm{d}x}\mathrm{d}x \\ &= p(x')n(x') - p(x)n(x) \end{aligned} \tag{6.1.10}$$

等式(6.1.10)表明基区的少数载流子(电子)电流取决于某一区域两端电子空穴乘积的差除以该区域多数载流子浓度的积分得到的值。如果把基区的两个空间电荷区边界作为该区域的两个边界,那么等式(6.1.10)中 $x=0$ 为积分下限,$x'=x_B$ 为积分上限。由等式(5.3.20)可知,边界上的 pn 乘积与结电压有关

$$p(0)n(0) = n_i^2 \exp\left(\frac{qV_{BE}}{kT}\right) \tag{6.1.11}$$

$$p(x_B)n(x_B) = n_i^2 \exp\left(\frac{qV_{BC}}{kT}\right)$$

这样,基区中 x 方向的电子电流可以表示成结电压 V_{BE} 和 V_{BC} 的函数

$$J_n = \frac{qn_i^2\left[\exp\left(\frac{qV_{BC}}{kT}\right) - \exp\left(\frac{qV_{BE}}{kT}\right)\right]}{\int_0^{x_B}\frac{p\mathrm{d}x}{D_n}} \tag{6.1.12}$$

在讨论等式(6.1.12)的含义之前,我们先稍做修改。首先,基区中的 D_n 与位置关系不大,可以用平均值 $\tilde{D}_n$ 替代,并且可以从等式(6.1.10)的分母中提到积分号外[2]。提出 $\tilde{D}_n$ 后,积分值为单位结面积下基区的多数载流子总量。与这个总量有关的电荷被称为 Q_B

$$Q_B = q\int_0^{x_B} p\mathrm{d}x \tag{6.1.13}$$

经过这些改动,从第一个结流向第二个结的电子电流密度可以表示成

$$J_n = J_S\left[\exp\left(\frac{qV_{BC}}{kT}\right) - \exp\left(\frac{qV_{BE}}{kT}\right)\right] \tag{6.1.14}$$

式中

$$J_S = \frac{q^2 n_i^2 \tilde{D}_n}{Q_B} \tag{6.1.15}$$

6.1.1 原型晶体管

我们推导出的描述晶体管原理的等式(6.1.12)与基区掺杂浓度分布无关,只与基区多子电荷的积分有关。从等式(6.1.12)推导出的等式(6.1.14)可以描述任意掺杂晶体管的工作过程。在更详细地研究普遍情形之前,我们先通过图6.1所示的原型晶体管来考虑晶体管的物理基础。

图6.2简单给出了原型晶体管在平衡和不同偏压条件下的能带图和电子(基区少数载流子)浓度分布。图6.2(a)表示了平衡(零偏压)条件下的晶体管能带图和电子浓度分布,其中基区只有少量电子,由于势垒的存在,电子无法从其他两个区域注入。两个结外加反向偏压时势垒增高,导致基区内的电子被耗尽,图6.2(b)是这种情况下的能带图与电子浓度分布。两个结上外加正向偏压时可以降低内建势垒,电子被注入到基区,导致基区电子浓度急剧升高,形成了两个结之间的电流,如图6.2(c)所示。

放大偏置 另一种偏压情况是一个结正偏,另一个结反偏。这种情况我们最感兴趣,因为这时晶体管有放大作用。图6.2(d)是这种放大偏置条件下的能带图和电子浓度分布。从图

---

[2] 不是直接空间平均值,而是 $\tilde{D}_n = \int_0^{x_B} p\mathrm{d}x \Big/ \int_0^{x_B} (p/D_n)\,\mathrm{d}x$。

中可以看出,正向偏压使结势垒降低,因此正偏结向基区注入电子。另一方面,基区中的电子一旦靠近反偏结,就会被拉入n型区。这样,电子(基区少数载流子)就从正偏结,也就是发射结,流到反偏的集电结。基区中集电结附近的电子在电场的作用下快速通过集电结空间电荷区进入n型集电区。

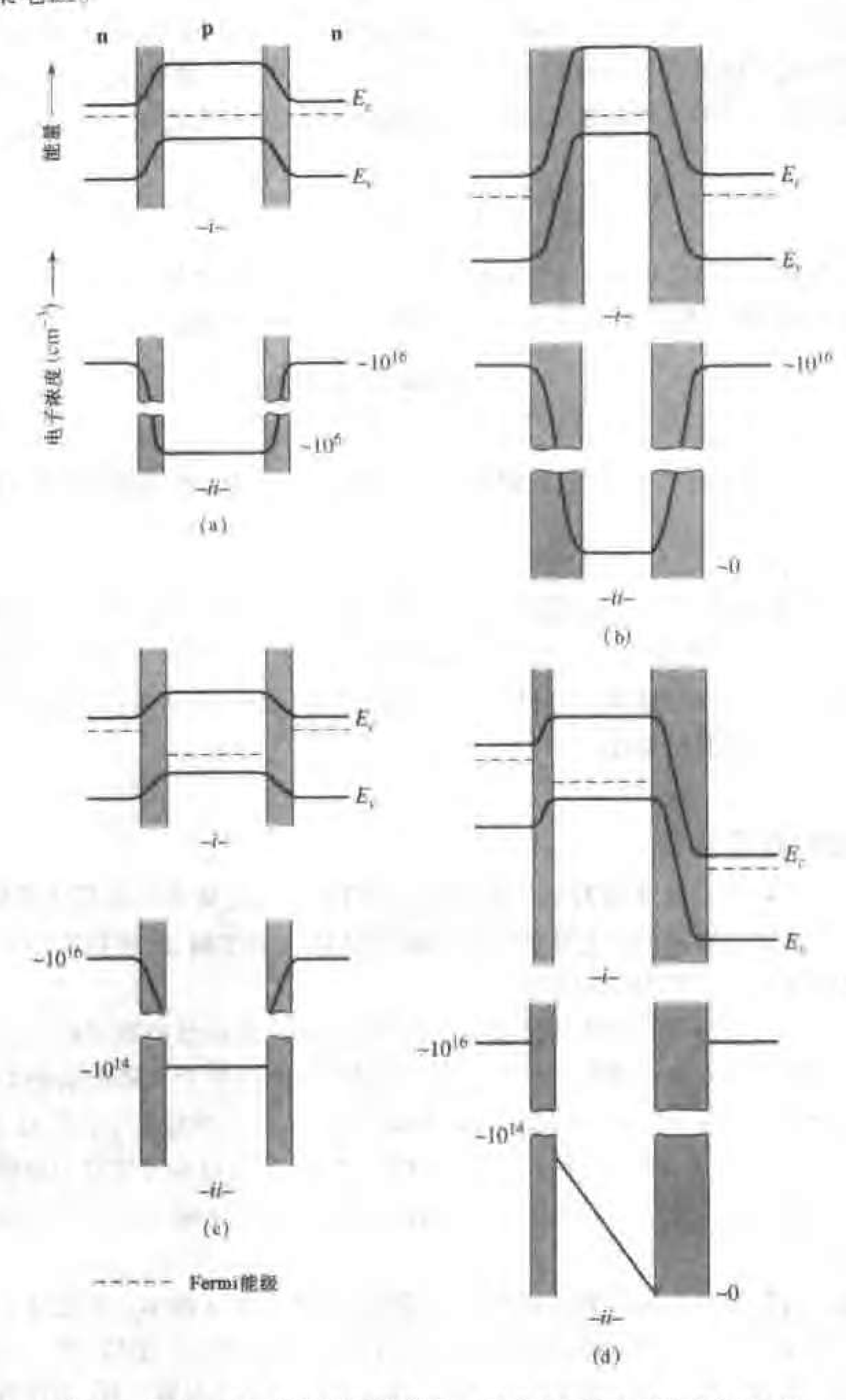


图6.2 图6.1所示晶体管的能带图。(*i*)和相应的电子浓度分布(*ii*)
(a)平衡状态;(b)两个结都反偏;(c)两个结都正偏;(d)一个结反偏,一个结正偏。阴影部分表示空间电荷区

为了分析放大偏置下的晶体管,并解释前面的一些定性讨论,我们简要考虑原型晶体管的放大或有源偏置工作区。图6.2(d)中的基区的少数载流子电子浓度为一条直线,这样的电子分布是根据5.3节中对短二极管的分析推导出来的,因为均匀掺杂晶体管基区电子分布与理想短二极管一样。也就是说,如果忽略基区复合作用,并且掺杂浓度为常数,那么,解连续性方程得到的电子浓度在x方向上线性变化。发射结边界的基区电子浓度随电压V_{BE}指数变化,集电结边界的基区电子浓度可以忽略。与等式(5.3.17)中的空穴类似,过剩电子浓度n'_p为线性分布。根据图6.1中的几何关系可以写出过剩电子浓度的表达式$n'_p=n_p-n_{po}$为

$$n'_p = n_{po}\left[e^{qV_{BE}/kT}\left(1-\frac{x}{x_B}\right)-1\right] \qquad 0 \leqslant x \leqslant x_B \tag{6.1.16}$$

一般在放大偏置条件下,V_{BE}远大于kT/q,基区中的电子电流主要是扩散电流。因为$J_n=qD_n(dn_p/dx)$,可以根据基区掺杂浓度N_a写出基区中的电子电流密度表达式

$$J_n = \frac{-qD_n n_i^2 \exp(qV_{BE}/kT)}{N_a x_B} \tag{6.1.17}$$

为了从等式(6.1.14)推出上述结果,我们从等式(6.1.13)推导出原型晶体管的Q_B为

$$Q_B = qN_a x_B \tag{6.1.18}$$

把这个Q_B值代入等式(6.1.15),忽略等式(6.1.14)中的$\exp(qV_{BC}/kT)$项,就可以得到电流密度表达式(6.1.17)。描述晶体管特性的方程组(等式(6.1.12)~(6.1.15))的重要性在于它们适用于晶体管的任意掺杂情况。这种分析方法最初用于优化晶体管掺杂分布[2],第7章将进一步介绍这种方法的其他应用。

6.1.2 集成电路晶体管

在推导晶体管放大偏置下的其他公式之前,我们先来研究集成电路晶体管的一般特性。首先比较图6.1中的简单的原型晶体管与集成电路晶体管的结构、原理和实际应用。集成电路晶体管的结构将在6.5节中详细介绍。

研究原型晶体管是因为可以不考虑额外的数学问题,而重点分析物理特性。原型晶体管中的掺杂分布可通过2.6节中提到的气相外延工艺获得[顺序生长不同掺杂的半导体材料经常被用来制作化合物半导体,有时也用来制作Ⅳ族半导体(例如在硅上生长锗硅合金)的异质结双极晶体管(HBT)]。6.6节将进一步分析HBT。然而,集成电路晶体管中硅的掺杂一般用第2章介绍的平面扩散方法得到,因此,集成电路晶体管的结构和掺杂分布与图6.1中描述的原型晶体管不同。

氧化物隔离的集成电路晶体管的俯视图和横截面图如图6.3所示。制造这个晶体管的工艺过程与5.5节介绍的氧化物隔离的集成电路二极管的制造过程很相似,惟一不同的是晶体管需要在p型区扩散形成一个高掺杂的n型区,即发射区。同二极管一样,晶体管的衬底为轻掺杂的p型材料。为了减小横向串联电阻,在集电结的下方需要增加一个重掺杂的n型埋层(图5.18)。6.5节将更详细地讨论晶体管中的掺杂分布、器件的几何结构和各参数之间的折衷问题,本节主要介绍器件的基本几何结构。

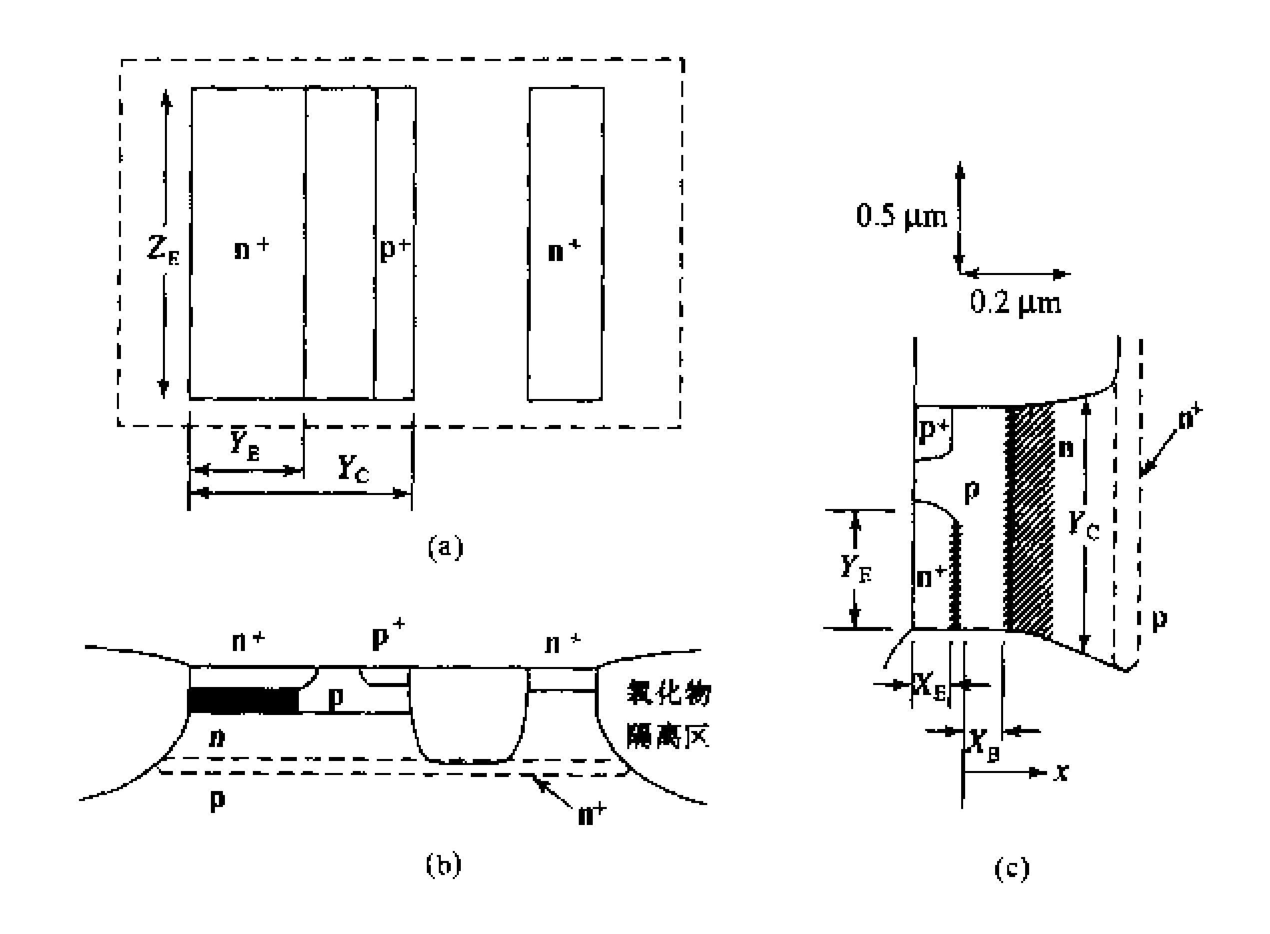


图6.3 氧化物隔离的npn集成电路晶体管的俯视图(a)和横截面图(b)和(c)。控制晶体管放大作用的区域为(b)中的阴影区,(c)为这个区域旋转90°后的放大图;x与y的比例关系只适用于图(c)

在晶体管设计中,准中性基区的宽度(图6.3(c)中的X_B)是最重要的参数之一。基区宽度一般为100nm量级(~10^{-5}cm)左右,高频晶体管更小。发射区宽度[③] Y_E在光刻允许的条件下尽可能小,大约几百纳米。晶体管的其他几何尺寸都很大。晶体管的放大作用主要决定于两个背靠背pn结间的距离,因此,在大多数偏置条件下,控制晶体管特性的主要区域为图6.3(b)中的阴影部分,图6.3(c)更清楚地指出了这个区域。图6.3(c)中交叉线阴影部分为发射结和集电结的空间电荷区。为了与图6.1中器件的放置方向一致,图6.3(c)被旋转了90°。

图6.3(c)表明晶体管放大作用主要发生在由准中性发射区宽度Y_E限定的区域,发射区一般是重掺杂,发射结空间电荷区基本上没有向发射区延伸,所以Y_E近似等于发射极条宽。尽管常规的集成电路工艺使得集电结宽度Y_C远大于发射结宽度Y_E,在很多情况下双极晶体管还是可以近似为一维器件。利用电流的一维近似,在定义了面积A后可以把电流密度J与电流I联系起来,由图6.3可知面积A为Y_E与Z_E的乘积。在进一步讨论结型晶体管理论时,我们将考虑图6.3中的其他几何尺寸,研究集成电路器件结构的限制。

6.2节将更详细地介绍放大偏置下的晶体管特性,其中的概念对理解晶体管的放大作用和开关作用非常重要。

6.2 放大偏置

V_{BE}正偏、V_{BC}零偏或者反偏的情况称做npn晶体管的放大偏置[④]。偏压导致电子在发射结

③ 注意:基区"宽度"方向垂直于表面,而发射区"宽度"方向平行于表面。

④ 具体地说,这种偏置对应于正向有源区;如果将V_{BE}与V_{BC}极性互换,则晶体管处于反向有源区。

注入,在集电结被收集。当 V_{BC} 为零或负,而 V_{BE} 远大于 kT/q 时,从等式(6.1.14)可求出从左向右流过集电结(图6.1中的 J_2)的电子电流

$$J_n \approx -J_S \exp\left(\frac{qV_{BE}}{kT}\right) \tag{6.2.1}$$

规定电流流入晶体管为电流的正方向,那么等式(6.2.1)中的 J_n 等于 $-J_C$[5]。从等式(6.2.1)可看出在放大偏置下,集电极电流随发射极-基极电压 V_{BE} 指数增大。

图6.4给出了对数坐标下集电极电流 I_C 与基极-发射极电压 V_{BE} 关系的实验曲线。几乎在整个电压区域里,集电极电流与基极-发射极偏压都呈指数关系。此外,如果将以10为底的对数坐标改成自然对数坐标,则集电极电流曲线的斜率则恰好等于 q/kT,这与等式(6.2.1)的结论相吻合。从等式(6.2.1)可以看出 $T=300\text{K}$ 时,基极-发射极电压每增加 $\Delta V_{BE}=(kT/q)\ln(10)=60\text{mV}$ 时,$|J_C/J_n|$ 上升10分贝。斜率偏离60mV/分贝是非理想情况,需要密切关注以便于指出工艺制造的问题。然而,一般来说集电极电流严格符合指数关系,因此保持基极-发射极偏置不变,根据集电极电流与温度的关系可以制出精密的温度计。

将 J_C 曲线的直线部分向下延伸至 $V_{BE}=0$ 处,在纵轴上得到的截距即为等式(6.2.1)中 J_S。一旦 J_S 已知,由等式(6.1.15)就可以得到准中性区的内建基区电荷

$$Q_{B0} = \frac{q^2 n_i^2 \tilde{D}_n}{J_S} \tag{6.2.2}$$

等式(6.2.2)中的其他所有参数都是已知的[6]。等式(6.2.2)中的 Q_{B0} 称为内建基区电荷,如等式(6.1.13)所示,Q_{B0} 代表发射结零偏时单位面积准中性基区的空穴电荷总量。这个电荷值由晶体管的工艺决定,可以从晶体管的电流电压关系得到。这种关系最早由H. K. Gummel提出[3],因此单位面积准中性基区的掺杂总量(每 cm^2)有时也被称为Gummel数 GN。

$$\int_0^{x_B} N_a(x)dx = \frac{Q_{B0}}{q} = \frac{qn_i^2\tilde{D}_n}{J_S} \tag{6.2.3}$$

等式(6.2.3)指出,在固定偏压下晶体管电流表达式中的 J_S 反比于Gummel数,也就是说,反比于基区掺杂总量。内建基区电荷 Q_{B0} 越少,固定偏压下的电流就越大,因此,设计原型晶体管时应尽可能降低基区掺杂浓度。但是,这样设计的主要缺点是即使在非常小的正向偏压下,小注入近似也可能无效(也就是说,$n_p(0)$ 可能接近于 N_a)。在第7章我们将了解到,晶体管在大注入时性能将下降。如果让基区的掺杂从发射结到集电结递减分布,则可避免这个问题,同时又能保证较小的 Q_{B0}。这种情况下,基区中注入少子浓度高的区域掺杂浓度高,注入少子浓度低的区域掺杂浓度低。幸运的是,集成电路扩散工艺可以实现晶体管基区的缓变掺杂,使得集成电路晶体管能具有这个优点。Kroemer还首次指出了缓变基区晶体管的其他优点[4],这些将在第7章中介绍。

晶体管制备过程中对 Q_{B0} 的控制是集成电路制造的关键。在Gummel数小于 10^{12}cm^{-2} 的高增益晶体管中,需要仔细控制 Q_{B0} 的值。

---

5 译者注:原文误为 $+J_C$

6 $\tilde{D}_n$ 可能是个特例,但是扩散系数随浓度变化不明显(图1.16),因此在解析计算中可以近似,数值模拟中可以用局部 D_n 值来避免近似。

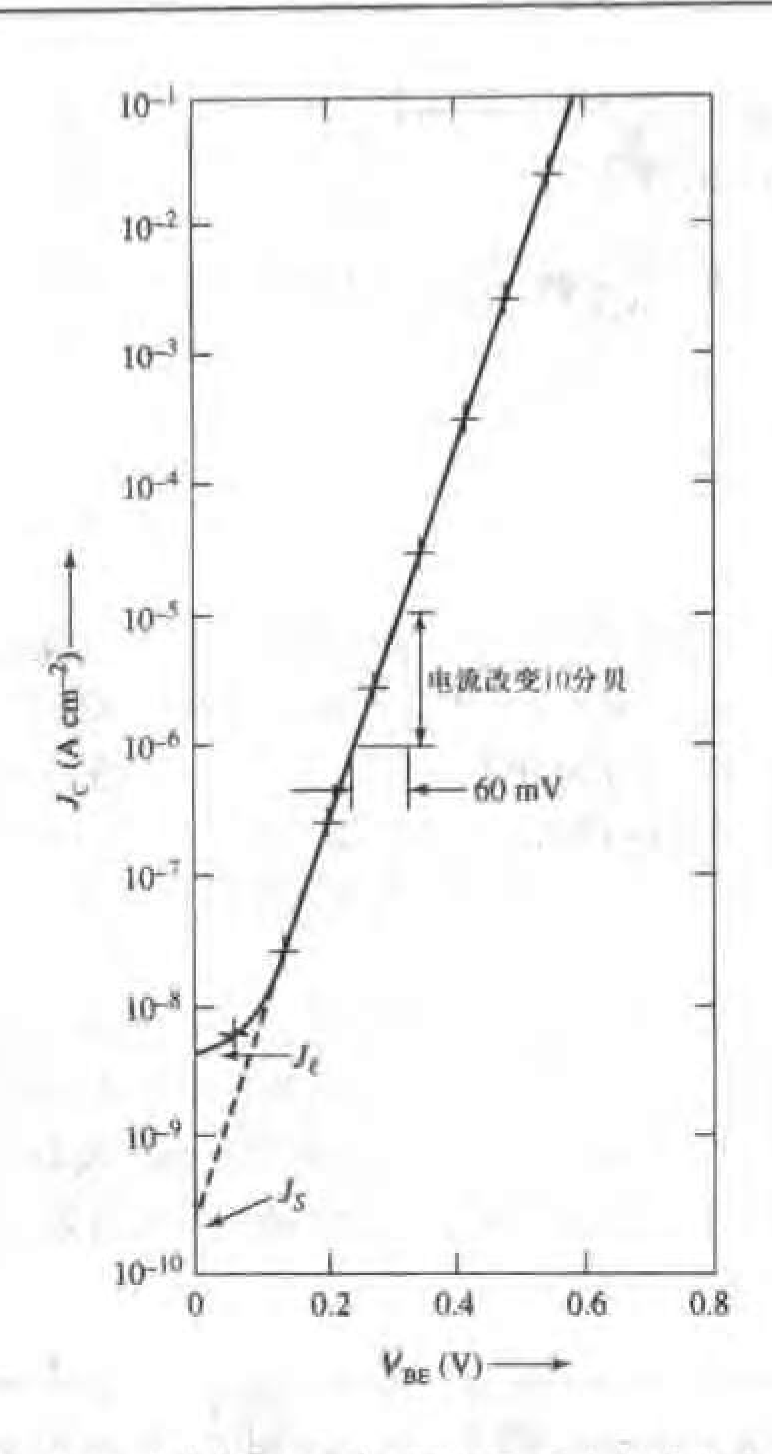


图6.4 温度为300K,放大偏置下,半对数坐标系中集成电路 npn 晶体管集电极电流与基极-发射极电压的关系曲线。十字表示数据点,曲线中直线部分的延长线指出了 J_S(等式(6.2.1));J_ℓ是漏电流部分

例题 计算 Gummel 数

假设图6.4中的原型 npn 晶体管基区宽度 $x_B = 0.5\mu m$,求晶体管的 Gummel 数(等式(6.2.3)),并计算发射结边界处的基区电子浓度为基区掺杂浓度的1%时所需的基极-发射极电压。

解:由图6.4可得 $J_S = 2.4 \times 10^{-10}$ A cm^{-2},由等式(6.2.3)可求出 Gummel 数 GN 为

$$GN = \frac{qn_i^2 \bar{D}_n}{J_S} = 1.4 \times 10^{11} \bar{D}_n$$

假设 $\bar{D}_n$ 为 20cm^{2}s^{-1}(见图1.16),则 $GN = 2.8 \times 10^{12}$cm^{-2},$N_a = \frac{GN}{x_B} = 5.6 \times 10^{16}$cm^{-3}。再由图1.16可知,当 $N_a = 5.6 \times 10^{16}$cm^{-3}时,$\bar{D}_n \approx 22$cm^{2}s^{-1},而不是假定的20cm^{2}s^{-1}。重新计算得到 $GN = 3.1 \times 10^{12}$cm^{-2}和 $N_a = 6.2 \times 10^{16}$cm^{-3},由这个结果查到的 $\bar{D}_n$ 与计算所用的值近似吻合。

因此晶体管的 Gummel 数为 $3.1\times10^{12}\text{cm}^{-2}$。

例题第二问所求的 V_{BE} 值满足下式

$$n'(x=0)=\frac{n_i^2}{N_a}\exp\left(\frac{qV_{BE}}{kT}\right)=0.01\times N_a=6.2\times10^{14}\ \text{cm}^{-3}$$

解得 $V_{BE}=0.67\text{V}$。

电流增益

迄今为止,对放大偏置下晶体管特性的讨论仅限于发射结与集电结之间的电子电流,相当于放大偏置晶体管的输出电流。注入基区的电子浓度随基极-发射极正向偏压指数变化,因此集电极电流是基极-发射极电压的指数函数(等式(6.2.1))。在放大偏置下,这些注入的电子大部分被集电结空间电荷区的电场收集,因此在发射结端可以控制集电极电流。任意固定的正向偏压 V_{BE} 下,基极电流越小,输入功率(V_{BE} 与基极电流的乘积)越小,晶体管的放大效率越高。

放大偏置晶体管的基极电流包括若干个产生机构,其中最直接的成分是与注入到准中性基区的电子复合的空穴电流。基极电流的第二个成分是发射结空间电荷区的复合电流。正向偏压下的发射结不仅从发射区向基区注入电子,同时基区也向发射区注入空穴,形成基极电流的第三个成分。图 6.5 标出了这些电流成分。为使晶体管能有效地放大,需要控制这些电流成分,使其远小于集电极电流。

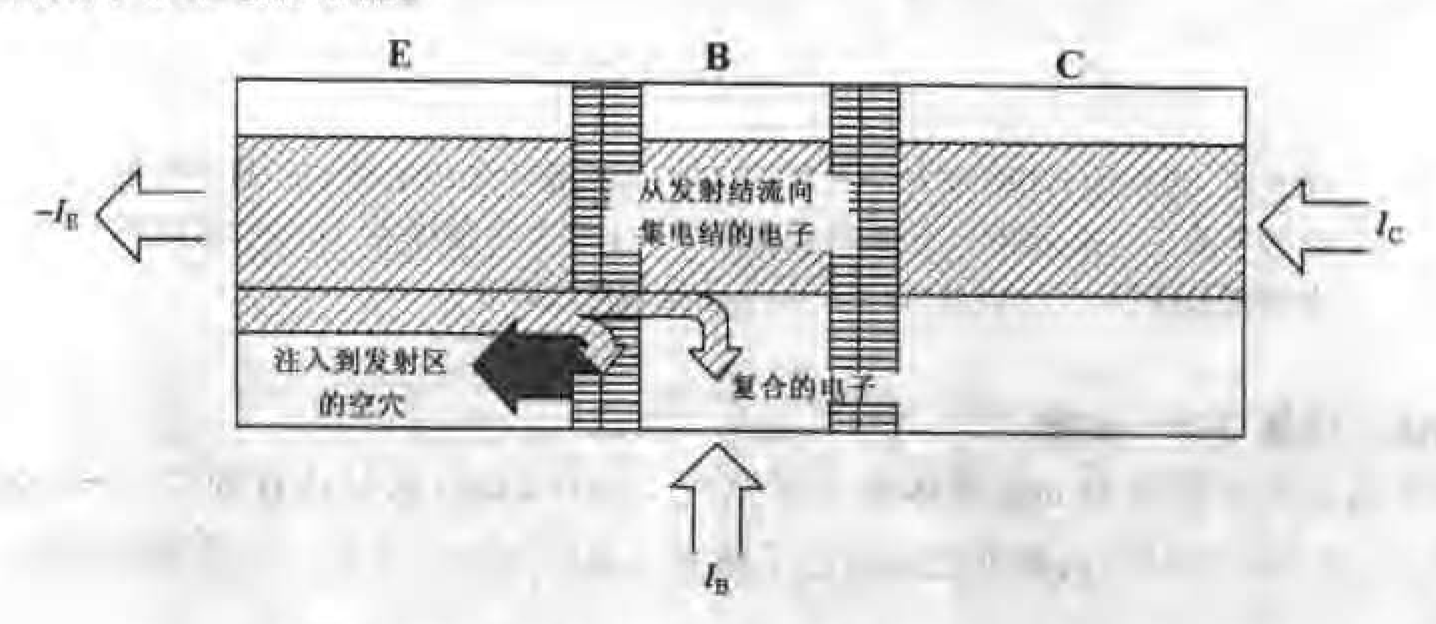


图 6.5 放大偏置晶体管的各电极电流与主要电流成分。没有标出集电结漏电流(图 6.4 中的 J_r)

在小注入条件下,大部分硅晶体管发射结空间电荷区复合电流小于其他电流成分(空间电荷区复合电流在化合物半导体晶体管中非常重要)。我们将在第 7 章研究影响晶体管性能的限制条件时单独讨论空间电荷区中的复合。

准中性基区中的复合可以用 5.3 节介绍的理论来分析。过剩少数载流子的复合与其浓度(这里 n' 为基区过剩电子浓度)成正比,因此基区总复合电流为

$$I_{rB}=qA_E\int_0^{x_B}\frac{[n-(n_i^2/N_a)]\mathrm{d}x}{\tau_n}\tag{6.2.4}$$

式中 $A_E = Y_E \times Z_E$ 为少数载流子的有效注入面积,如图 6.3 所示。在给定的偏置条件下为减小 I_{rB},需要增大基区寿命 τ_n,同时减小基区宽度 x_B。

放大偏置下大部分基区中注入的过剩电子浓度 n' 远大于平衡态电子浓度(n_i^2/N_a),并且寿命与 x 关系不大,等式(6.2.4)可以简化为

$$I_{rB} = \frac{qA_E}{\tau_n}\int_0^{x_B} n'\mathrm{d}x \tag{6.2.5}$$

对于基区均匀掺杂的晶体管,例如图 6.1 所示的原型晶体管,n' 随 x 近似线性变化,这样等式(6.2.5)中的积分很容易求出

$$I_{rB} = \frac{qA_E n_i^2 x_B}{2N_a\tau_n}\left[\exp\left(\frac{qV_{BE}}{kT}\right) - 1\right] \tag{6.2.6}$$

虽然等式(6.2.6)只适用于特殊情况,但是通常情况下基区复合电流与[exp(qV_{BE}/kT) - 1]的比例关系依然成立。

基区输运系数用来表征载流子在基区内复合所引起的损失,通常用 α_T 表示,即

$$\alpha_T = \frac{|I_{nE}| - |I_{rB}|}{|I_{nE}|} = 1 - \left|\frac{I_{rB}}{I_{nE}}\right| \tag{6.2.7}$$

式中 I_{nE} 代表从发射结注入的电子电流。对于基区均匀掺杂的晶体管,由等式(6.1.17)与(6.2.6)可得

$$\alpha_T = 1 - \frac{x_B^2}{2D_n\tau_n} = 1 - \frac{x_B^2}{2L_n^2} \tag{6.2.8}$$

等式(6.2.8)不能直接用于分析集成电路晶体管,因为基区的掺杂不均匀,但有时也可用该式近似分析这类器件。扩散工艺形成的缓变基区可以提高 J_S,使得输出电流固定时需要注入的少子减少,因此缓变基区可以优化输运系数,由等式(6.2.8)计算出的 α_T 可以被认为是"最坏情况"下的值。假设少数载流子的扩散长度为 10μm,x_B 为 0.3μm,则 $\alpha_T = 0.9996$。高性能集成电路晶体管的基区宽度很窄,准中性基区内的少数载流子复合引起的电流损失也很小。

在大多数集成电路晶体管中,基区向发射区注入的多数载流子(空穴)是基极电流的主要成分,此时的发射结相当于一个正偏二极管,5.3 节已给出了这个电流的表达式。对于一个给定的器件,判定注入到发射区的空穴在哪里复合是非常重要的,这样才能写出空穴电流的正确表达式。首先考虑图 6.1 中的原型晶体管,发射极与发射结空间电荷区边缘的距离为 x_E。如果 x_E 远大于发射区空穴扩散长度,注入的所有空穴在到达欧姆接触电极之前就已经复合,因此,由 5.3 节知,发射区中的过剩空穴将以指数形式下降,如图 6.6(a)所示,空穴电流正比于准中性发射区边缘的空穴浓度梯度。

$$I_{pE} = \frac{-qA_E n_i^2 D_{pE}}{N_{dE}L_{pE}}(e^{qV_{BE}/kT} - 1) \tag{6.2.9}$$

如果发射极离基区很近($x_E \ll L_{pE}$),则空穴浓度是 x 的线性函数(见图 6.6(b)),空穴电流为

$$I_{pE}' = \frac{-qA_E n_i^2 D_{pE}}{N_{dE}x_E}(e^{qV_{BE}/kT} - 1) \tag{6.2.10}$$

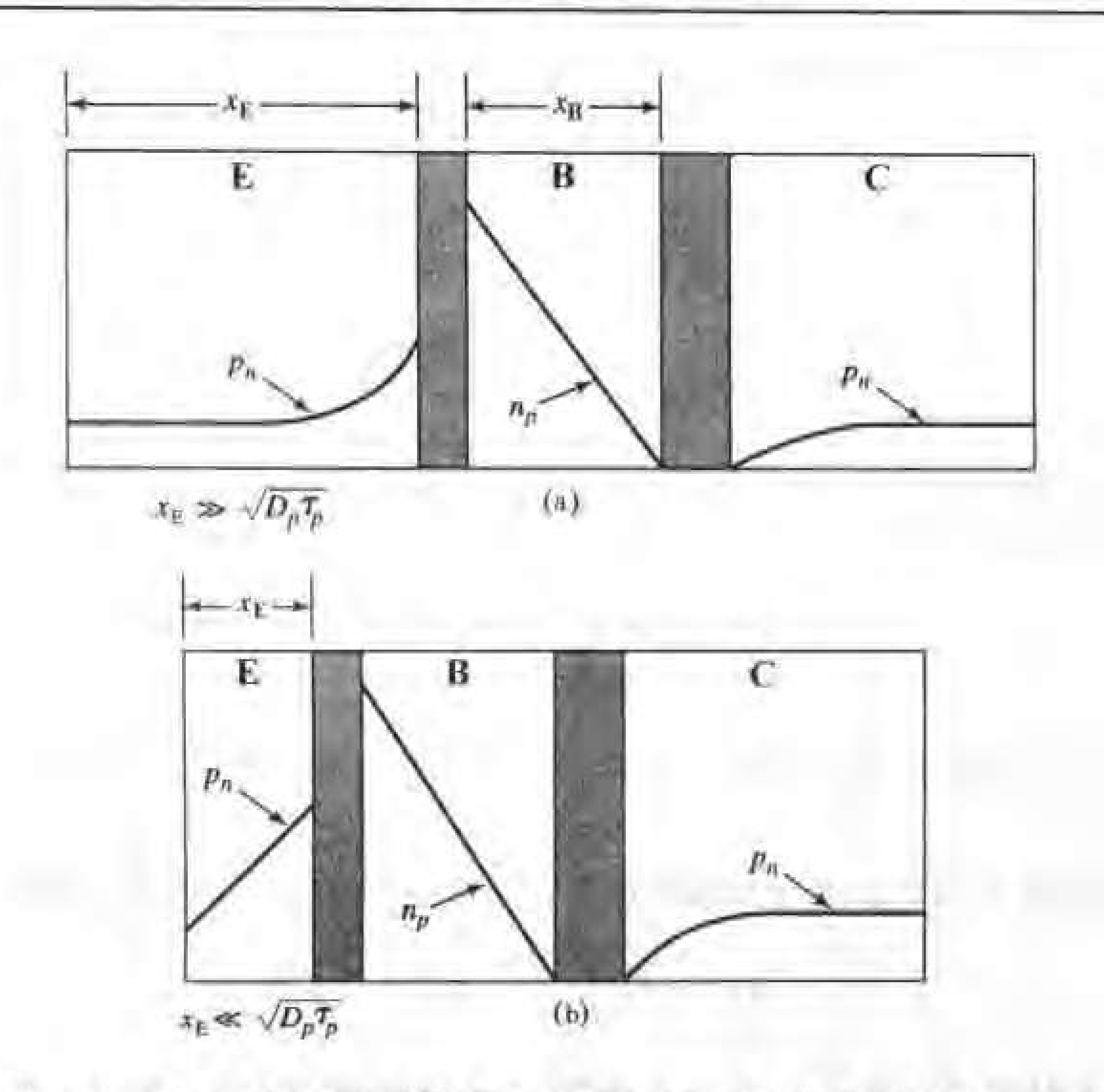


图6.6　图6.1中的原型晶体管在放大偏置下的少数载流子分布。(a) x_E远大于发射区空穴扩散长度;(b) x_E远小于发射区空穴扩散长度

大多数集成电路晶体管的发射极到基区的距离小于一个扩散长度。因为发射区掺杂不均匀，所以注入到发射区的空穴会受到内建电场的影响。公式推导与6.1节推导电子在基区内的输运类似，但是此时多子(电子)电流非零。

发射结注入:非均匀掺杂　下面讨论图6.3中的集成电路晶体管的非均匀掺杂情况。在热平衡状态下，杂质的分布会在发射区形成内建电场。然而由4.1节知除了空间电荷区之外，大部分的非均匀掺杂区域能够近似为中性或准中性。这个近似意味着 $n_0(x)\approx N_d(x)$，此时由等式(4.1.13)可求出热平衡时的内建电场

$$\mathscr{E}_0(x) = -\frac{kT}{q}\frac{1}{N_{dE}(x)}\frac{dN_{dE}(x)}{dx} \tag{6.2.11}$$

我们预测发射结正偏时准中性发射区内的电场与热平衡状态相比改变不大。在一阶分析中用 $\mathscr{E}_0$ 加上由 $V_{BE}\neq 0$ 导致的 $\mathscr{E}_a$ 来表示该偏置下的电场，这个假设与小注入条件相符(即 $p_n' = n_n' \ll n_0$)。这样发射区中的电子和空穴电流可以写为

$$J_n = q\mu_n(n_{no} + p_n')(\mathscr{E}_0 + \mathscr{E}_a) + qD_n\left(\frac{dn_{no}}{dx} + \frac{dp_n'}{dx}\right)$$

和

$$J_p = q\mu_p(p_{no} + p_n')(\mathscr{E}_0 + \mathscr{E}_a) - qD_p\left(\frac{dp_{no}}{dx} + \frac{dp_n'}{dx}\right) \tag{6.2.12}$$

平衡状态下漂移和扩散作用能相互抵消($\mathscr{E} = \mathscr{E}_0$时)，等式(6.2.12)可以写成

$$J_n = q\mu_n n_{no}\mathscr{E}_a + qD_n\frac{dP_n'}{dx} \tag{6.2.13}$$

和

$$J_p = q\mu_p(p_{no} + p_n')\mathscr{E}_a + q\mu_p p_n'\mathscr{E}_0 - qD_p\frac{dP_n'}{dx} \tag{6.2.14}$$

式中 $p_n'(\mathscr{E}_a+\mathscr{E}_0)$ 与 $n_{n0}\mathscr{E}_a$ 相比可以忽略。

等式(6.2.13)中的第一项代表多数载流子电子的漂移流。与理想的二极管相同,很小的电场 $\mathscr{E}_a$ 就能提供这一漂移电流,所以一般 $\mathscr{E}_a$ 小于 $\mathscr{E}_0$。将这一假设用于空穴电流密度表达式(6.2.14)中,发现第二项的值远大于第一项,因此,

$$J_p = q\mu_p p_n'\mathscr{E}_0 - qD_p\frac{dP_n'}{dx} \tag{6.2.15}$$

将热平衡状态下的电场表达式(6.2.11)代入上式,得到与基区电子电流表达式(6.1.4)相似的形式

$$J_p = -\frac{qD_p}{N_d(x)}\frac{d}{dx}[p_n'(x)N_d(x)] \tag{6.2.16}$$

下面考虑 $x_E \ll L_{pE}$ 的短发射区情况,这时n型区内的空穴复合可以忽略,空穴电流不再是位置的函数,将等式(6.2.16)在负 x 的方向上从发射区任意一点 x 到 $p_n'=0$ 的欧姆接触处积分,得到

$$J_p\int_x^{x_E}\frac{N_d(x')dx'}{D_p} = qp_n'(x)N_d(x) \tag{6.2.17}$$

在发射结空间电荷区边缘 $-x_n$ 处,将等式(5.3.8)得到的少数载流子浓度代入等式(6.2.17),得到发射区的空穴电流为

$$I_{pE}'' = -\frac{q\tilde{D}_p n_i^2 A_E(e^{qV_{BE}/kT}-1)}{\int_{-x_n}^{x_E}N_d dx} \tag{6.2.18}$$

如果发射区均匀掺杂,等式(6.2.18)可以化简成等式(6.2.10)。

如果准中性区中的复合不能忽略,部分载流子在到达发射极之前就已经复合,这时等式(6.2.16)仍然正确,但是不能将 J_p 移出到积分号外得到简单的表达式(6.2.18),空穴电流要比等式(6.2.18)求出的结果大。这种情形可以进一步近似,这里就不再讨论了。

以上两种情形下的 I_{pE}(等式(6.2.10)和(6.2.18))主要由表面复合决定,A_E 的值不能简单确定,因为欧姆接触界面的表面复合远大于氧化层界面。要得到准确的 I_{pE},必须用接触孔面积代替等式(6.2.10)与(6.2.18)中的结面积 A_E,6.5节在讨论平面扩散晶体管时将进一步分析这个问题。

发射区向基区注入电子的效率用发射效率来衡量,通常用符号 γ 表示

$$\gamma = \frac{|I_{nE}|}{|I_{nE}|+|I_{pE}|} = \frac{1}{1+|I_{pE}/I_{nE}|} \tag{6.2.19}$$

总发射极电流 I_E 等于($|I_{nE}|+|I_{pE}|$),流过发射结的电子电流 I_{nE} 可以表示成 γI_E,因此可以用推导 α_T 的方法得到集成电路晶体管中 γ 的近似表达式。利用原型晶体管(图6.1)的简单理论,以及合适的集成电路晶体管的尺寸参数和平均掺杂浓度,将等式(6.1.12)和(6.2.10)代入等式(6.2.19)得到

$$\gamma \approx \frac{1}{1+\dfrac{x_B N_{aB}\tilde{D}_{pE}}{x_E N_{dE}\tilde{D}_{nB}}}=\frac{1}{1+\dfrac{GN_B\,\tilde{D}_{pE}}{GN_E\,\tilde{D}_{nB}}} \tag{6.2.20}$$

等式的第二项引入了发射区 Gummel 数,其表达式与基区类似(等式(6.2.3))。

等式(6.2.20)经常用来表示双极晶体管的发射效率,但是不能直接用于集成电路晶体管,因为等式的推导过程中没有考虑发射区重掺杂引起的两种效应。第一种效应是1.1节(图1.3和等式(1.1.33))提到的当硅中的掺杂浓度高于 $10^{18}\,cm^{-3}$ 时,会出现禁带变窄和相应的本征载流子浓度 n_i 升高。由于注入 npn 晶体管发射区的空穴浓度正比于 n_i^2(等式(6.2.10)),重掺杂发射区的禁带变窄将引起少数载流子的注入增加,发射效率降低。第二种效应是 Auger 复合(等式(5.2.16))引起的寿命降低,这是由于多数载流子的浓度较高。正如等式(6.2.10)所示,发射区少子寿命的降低导致复合不能忽略。

由于这两种效应的存在,反向注入(空穴注入到发射区)电流表达式(6.2.10)的推导不适用于发射区重掺杂的情况。这里不对重掺杂引起的效应作严格的分析,而只是减小等式(6.2.20)中发射区的 Gummel 数。Gummel 数的减小幅度相当大,如果掺杂浓度接近 $10^{21}\,cm^{-3}$,有效发射区 Gummel 数只占发射区总掺杂浓度的几个百分点。

例题　BJT 的发射效率

利用等式(6.2.20)计算集成电路晶体管的发射效率。已知晶体管发射区空间电荷区边界距离电极 $0.8\,\mu m$,基区 Gummel 数 GN_B 为 $3\times10^{12}\,cm^{-2}$,假设发射区表面 $x=0$ 处掺杂浓度为 $6\times10^{20}\,cm^{-3}$,在空间电荷区边界 $x=0.8\,\mu m$ 处指数下降至 $5\times10^{16}\,cm^{-3}$。重掺杂效应使发射区 Gummel 数(GN_E)降至实际总掺杂浓度的2%。

解:首先需要求出发射区 Gummel 数,并估算比值 $\tilde{D}_{pE}/\tilde{D}_{nB}$。假设在发射区掺杂浓度随特征长度 λ 指数变化

$$N_{dE}=N_{dE0}\exp\left(-\frac{x}{\lambda}\right)$$

分别代入在 $x=0$ 与 $x=0.8\,\mu m$ 处的值,可求出 λ

$$\lambda^{-1}=\frac{\ln(1.2\times10^4)}{0.8\times10^{-4}}\quad 或\quad \lambda=85.2\ \text{nm}$$

因此单位面积发射区的总掺杂浓度为

$$\int_0^{x_E} N_{dE}(x)\mathrm{d}x \approx N_{dE0}\lambda = 5.1\times10^{15}\ \text{cm}^{-2}$$

有效发射区 Gummel 数是单位面积发射区总掺杂浓度的2%,即 $1.02\times10^{14}\,cm^{-2}$。用有效发射区 Gummel 数除以发射结深度得到平均发射区掺杂浓度 $N_{davg}=1.28\times10^{18}\,cm^{-3}$,利用

这个值查图 1.16 可得到发射区扩散系数 $\tilde{D}_{pE} \approx 4.0\text{cm}^2\text{s}^{-1}$，从 6.2 节的例题知基区扩散系数 $\tilde{D}_{nB}$ 约为 $22\text{cm}^2\text{s}^{-1}$，将这些参数代入等式(6.2.20)中可以得到

$$\gamma = \frac{1}{1 + \dfrac{3 \times 10^{12} \times 4.0}{1.02 \times 10^{14} \times 22}} = 0.9947$$

发射效率很高，但是没有等式(6.2.8)后面的举例说明求出的基区输运系数 α_T(0.9996)高。这个结果很典型，集成电路晶体管中的基区输运系数比发射效率更接近于 1。

本例题利用了几个近似条件，其中用平均掺杂浓度对应的扩散系数来代表分布参数 D 至少看上去是合理的，然而，为了考虑到禁带变窄和 Auger 复合效应而假设发射区 Gummel 数下降为原来的 1/50 似乎有些牵强。我们可以用数值模拟的方法对这个例题进行更严格的求解。

正向放大偏置下集电极电流 I_C 与发射极电流 I_E 的比值用 α_F 表示，分析可知 α_F 为发射效率 γ 与基区输运系数 α_T 的乘积

$$\alpha_F = \gamma\alpha_T \tag{6.2.21}$$

根据基尔霍夫定律，流入晶体管的总电流等于零，那么

$$I_B + I_E + I_C = 0$$

$$I_B - \frac{I_C}{\alpha_F} + I_C = 0$$

或

$$I_C = \frac{\alpha_F I_B}{(1 - \alpha_F)} = \beta_F I_B \tag{6.2.22}$$

式中 $\beta_F \equiv I_C/I_B$ 为共发射极电流增益[7]。因为 α_F 接近于 1，所以 β_F 的值很大(一般为 100 左右)。例如上面的例题，$\alpha_T = 0.9996$ 与 $\gamma = 0.9947$ 可以求出 $\alpha_F = 0.9943$ 与 $\beta_F = 174$。α_F 的微小变化，例如晶体管制造工艺的波动，将引起 β_F [$d\beta_F = d\alpha_F/(1-\alpha_F)^2$] 较大的改变，因此 β_F 很难被精确控制。电路设计者只能保证 β_F 的值足够大，在给定的工艺下 β_F 的值也可能会有很大的不同。

由上面的例题可知在双极晶体管中发射效率 γ 是限制共发射极电流增益 β_F 大小的主要因素，而且，通过增大发射区 Gummel 数来提高 γ 的方法受到禁带变窄和 Auger 复合的限制。

分析等式(6.2.9)可知，基区向发射区注入空穴限制了晶体管的电流增益，而该注入与准中性发射区边缘的空穴浓度梯度成正比。如图 6.6(b)所示，当 $x_E \ll L_p$(空穴扩散长度)时，发射区的空穴浓度分布与发射区表面处的浓度有关。对于典型的发射极金属接触，电极处过剩空穴浓度为零，使发射区的空穴浓度梯度达到最大值，限制了发射效率和晶体管增益。如果可以提高单晶晶体管发射极表面的空穴浓度，降低空穴浓度梯度，则可降低基区向发射区注入的空穴电流，晶体管增益将显著提高。

可以通过在金属接触与重掺杂单晶发射区之间增加一层 n^+ 多晶硅(多晶硅见 2.6 节)来

---

[7] 直流电流增益经常用 h_{FE} 表示。

提高表面空穴浓度。单晶硅-多晶硅界面处的空穴浓度可以很高，因此提高了发射效率，同时也提高了晶体管增益 β_F。

人们已经提出了很多物理机构来解释这个改进，但主要机理不是很确定，而且可能与具体的工艺流程有关。一种解释认为在单晶硅和多晶硅的界面处存在一层很薄的势垒（≤1nm），可能是残留的二氧化硅所致，它能够阻止空穴从单晶发射区向多晶硅流动，却对电子从多晶硅流入单晶发射区影响不大。[这种理论的基础是电子要越过的势垒低于空穴势垒（$E_{c,oxide} - E_{c,Si} < E_{v,Si} - E_{v,oxide}$）]。另一种解释依赖于多晶硅（有大量的不完整的晶体结构）中的少数载流子空穴的扩散长度小于单晶硅。这两种理论的结论都是单晶发射区中的空穴浓度梯度下降，减少了基区向发射区注入的空穴，提高了发射效率。这种"多晶硅发射极"双极晶体管的结构将在6.5节中介绍。

这一节所涉及的理论只适用于直流偏置条件，推导出的方程适用于所有电流范围。在放大电路中，需要相应的方程来反映电压围绕直流偏置点改变时的响应，第7章将考虑晶体管小信号特性，并推导出相应的等效电路以方便电路设计，同时还将研究晶体管的频率响应。

迄今为止，我们只考虑了直流和低频情形，另外需要指出的是，考虑发射区非均匀掺杂对注入空穴电流的分析同样适用于集成电路pn结，例如，方程6.2.18给出的注入空穴电流，与5.6节扩散pn结二极管（见图5.19(b)）中注入p型区的电子电流形式类似。

6.3　晶体管开关工作

6.1节讨论晶体管特性时，强调了注入到基区的电子（基区少数载流子）的重要性。同样，我们可以通过研究基区电子的注入、存储、抽取和输运来理解晶体管开关特性，并推导出应用广泛的晶体管模型。

工作区

图6.7根据外加结电压结 V_{BE} 和 V_{BC} 给出了器件的工作区，这有助于理解晶体管的开关工作。第四象限（V_{BE} 为正而 V_{BC} 为负）对应6.2节讨论的正向放大区。

反向放大偏置　在第二象限，V_{BC} 与 V_{BE} 的极性与正向放大偏置刚好相反，此时晶体管处于反向放大区。在这个工作区，npn晶体管在集电结处发射电子，在发射结处收集电子，此时晶体管各参数可以与正向放大偏置下的晶体管参数一一对应，通过给参数加上下标 R 来表示反向测量值。例如在反向放大偏置下电流在发射极输出，输出电流与输入电流（基极电流）的比值可以表示为 $\beta_R = I_E/I_B$。图6.1中的原型晶体管在反向放大偏置下的少数载流子分布如图6.8所示，注意与图6.6(a)比较。

图6.1所示的原型晶体管，其器件结构是对称的，发射区与集电区的掺杂浓度相同，并且两个结的面积相等，因此正向放大偏置和反向放大偏置的差别不大。

与原型晶体管不同，集成电路晶体管（见图6.3）的几何结构和掺杂浓度都不对称，因此反向放大偏置下的少数载流子分布与正向放大偏置有明显不同。首先，由于掺杂浓度比不同，反向放大偏置模式下的注入效率远低于正向放大区；其次，基区内建电场的方向阻碍从集电结注入的电子向发射结方向移动；第三，注入结面积远大于收集结面积，从集电结注入的电子在基

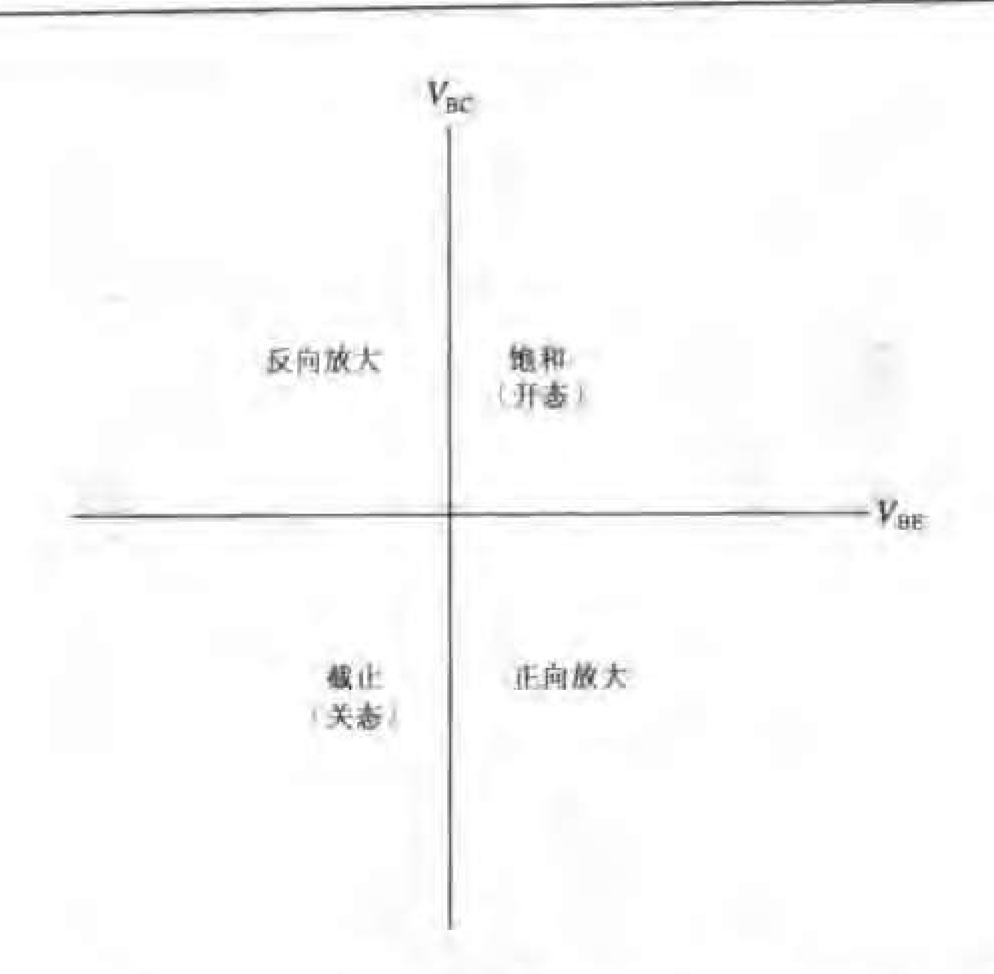


图6.7 由基极-发射极和基极-集电极偏压定义的npn晶体管的工作区

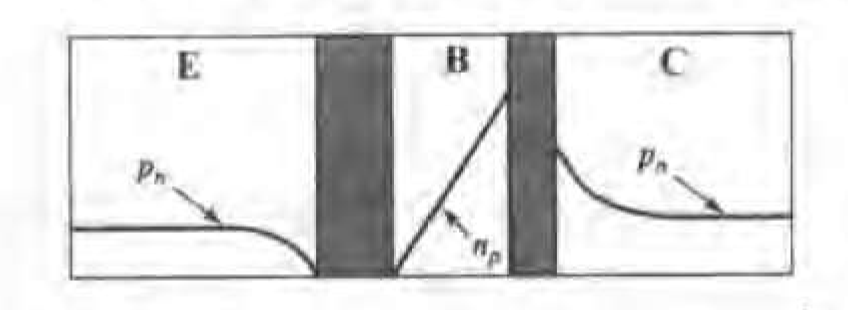


图6.8 图6.1所示的原型晶体管在反向放大偏置下的少数载流子浓度分布(与图6.6(a)比较)

极电极和钝化氧化层界面处有大量的损失，而正向放大区这类损失不大。这些不对称条件使得反向放大偏置下的增益低于正向放大偏置，这一点有很重要的实际意义，我们将在第7章继续讨论。

饱和 图6.7的第一象限中，基极-发射极和基极-集电极偏压都为正，这种偏置条件称为饱和[8]。开关晶体管的这种偏置状态称为"开启"状态。饱和晶体管的两个结都向基区注入电子，基区电子浓度显著升高，如图6.9所示。将图6.9中的基区少数载流子分布与图6.6和图6.8比较，我们发现饱和区可以看做是正向放大区和反向放大区的叠加，其物理机制是两个结同时注入和收集电子。向基区注入电子是因为pn结上的电势降低，从基区抽取电子是因为pn结上的电场能够将电子扫出基区。这个叠加关系与晶体管的几何结构无关，在集成电路晶体管和原型晶体管中都存在。

存储在饱和晶体管基区的大量电荷必须在晶体管"关断"前被抽走，这将降低晶体管的开关速度。3.6节讨论的Schottky二极管的箝位效应可以防止晶体管进入饱和区，从而提高开关速度。

---

[8] 集电极电流主要决定于外部电路条件，而不是晶体管本身，因此使用术语"饱和"。

截止 当V_{BE}与V_{BC}同时为负或零的时候，晶体管处于截止状态，此时基区内的电子数目小于或等于平衡状态下的电子数目，因此只有很小的电流在发射极和集电极之间流过。当两个结都为负偏压时，基区原本就很少的"内建"少数载流子大部分被耗尽，此时晶体管的直流行为接近于开路。原型晶体管截止状态的少数载流子浓度分布如图6.10所示。

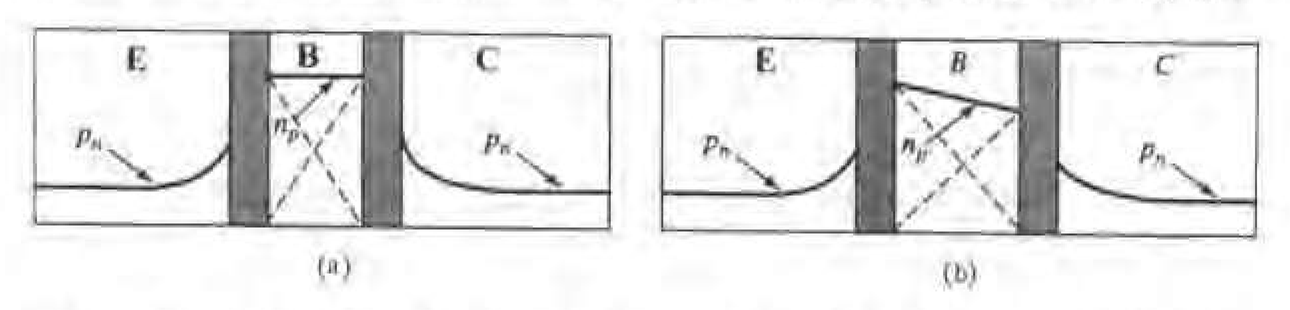


图6.9 图6.1所示的原型晶体管在饱和状态下少数载流子浓度分布;(a)集电极与发射极之间无电流的情况,(b)集电极与发射极之间有电流的情况

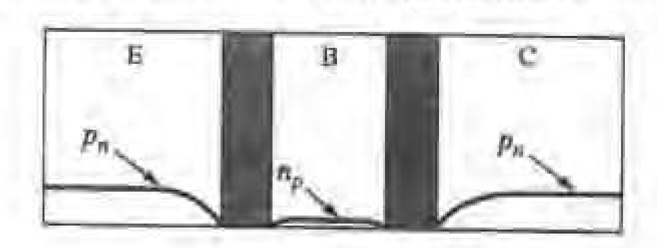


图6.10 图6.1所示的原型晶体管在截止状态下少数载流子浓度分布

尽管截止状态的晶体管中几乎没有直流电流，但它的行为并不完全等同于开路，而是类似4.3节描述的反偏pn结。我们可以利用数值模拟或者第4章提到的突变结和线性缓变结理论得出存储在结中的电荷Q_V随电压变化的关系。5.4节的例题给出了突变pn结二极管的存储电荷与外加偏压的关系，图6.11为类似的突变结和线性缓变结的Q_V曲线。下面的例题和习题6.7考虑了曲线的推导。图6.11中的小图明确表明，Q_V代表的存储电荷是与热平衡状态相比空间电荷区电荷的改变。

为使晶体管处于截止状态，开关电源必须向结提供增加的存储电荷，因此，截止晶体管特性类似电容，尽管电容(dQ/dV)不是常数。

例题 存储电荷与晶体管开关

类似图6.1中所示的npn原型晶体管，已知基区掺杂浓度$N_a = 10^{16}\text{cm}^{-3}$，发射区和集电区掺杂浓度为$N_d = 10^{19}\text{cm}^{-3}$，结面积为$10^{-5}\text{cm}^2$，晶体管处于截止状态，$V_C = 3\text{V}$，$V_E = 0\text{V}$，$V_B = -3\text{V}$。要想使基极电压$V_B = 0\text{V}$，则需要向基区提供多少电荷？

解：由等式(4.2.10)可求出发射结和集电结的$\phi_i = 0.872\text{V}$，由等式(4.3.1)求出图6.11中的K_A为

$$K_A = \sqrt{2\epsilon_s q N_a \phi_i} = 53.9\text{nC} \quad \text{cm}^{-2}$$

已知结面积为10^{-5}cm^2，那么$K_A \times A = 0.54\text{pC}$可作为$Q_V$表达式中的系数。起始状态集电结电压$V_a = -6\text{V} = -6.88 \times \phi_i$，由图6.11可知存储在集电结的电荷$Q_{VC} = -1.8 \times K_A A = -0.972\text{pC}$。当$V_B = 0\text{V}$时，$V_a = -3\text{V} = -3.44 \times \phi_i$，$Q_{VC} = -0.59\text{pC}$，因此基区提供的电荷为0.38pC。发射结的偏压从-3V变化到0V，存储的电荷由-0.59pC变化到0pC。

基极提供的总电荷是这些电荷的总和，即0.97pC。假设基极电源可以提供最大值为1mA的电流，那么本例中的电压转换时间为0.97ns。这类计算经常用于晶体管开关电路设计。

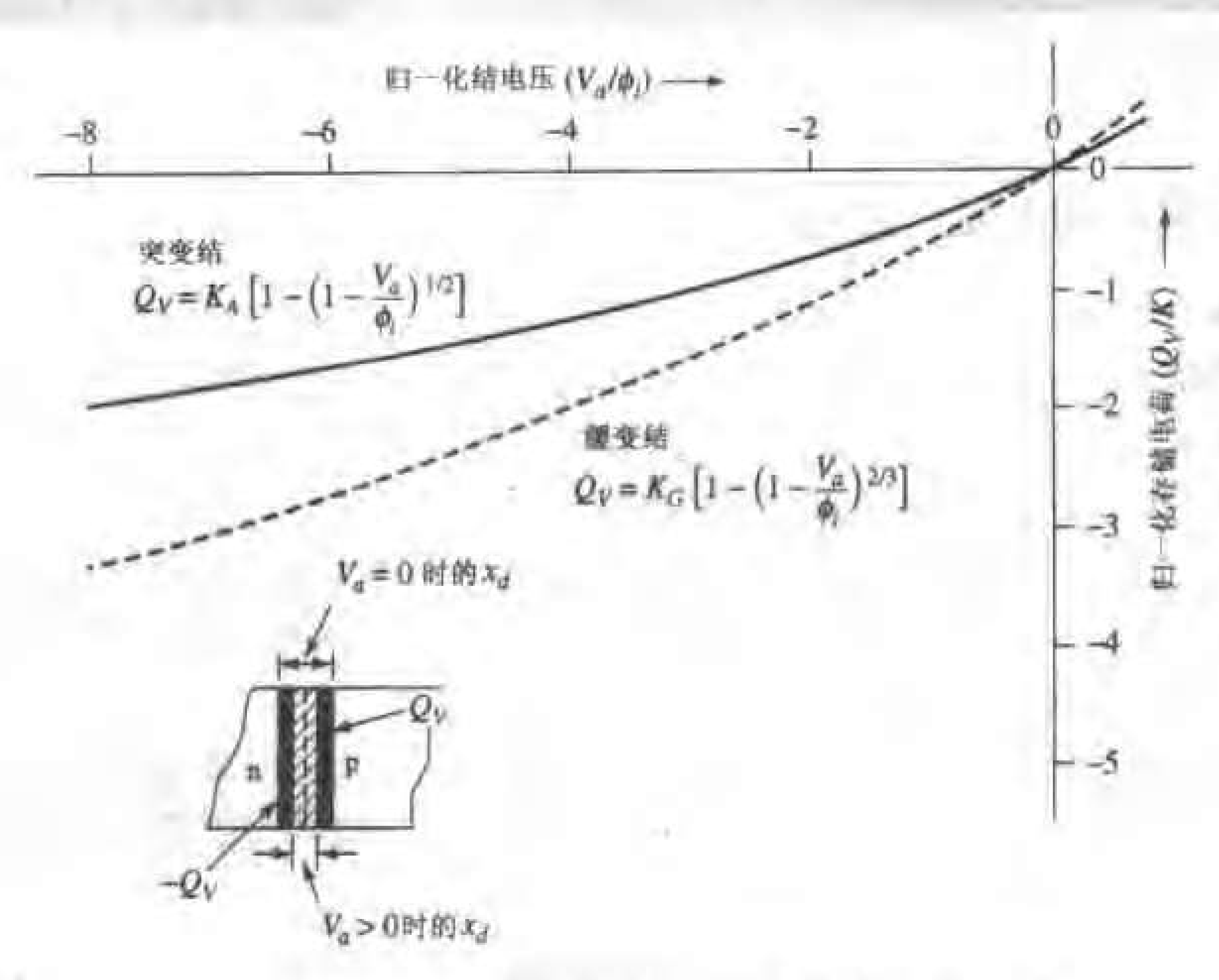


图6.11　突变结和线性缓变结的空间电荷区中，归一化存储电荷 Q_V/K 随归一化反偏结偏压 V_a/ϕ_i 变化的曲线

从图6.4的放大偏置曲线中可得到晶体管截止时流过的小电流。从图中我们注意到，当集电结反偏，V_{BE}降低至接近于零时，集电极电流将偏离等式(6.2.1)预示的指数关系，并且 J_c 趋向于一个与 V_{BE} 无关的值 J_ℓ。当基极-发射极电压很小时，发射结注入的电子数目很小，两个结之间的晶体管放大作用可忽略。集电结在反向偏压情况下接近于一个独立的pn结二极管，因此，晶体管此时会有一个很小的二极管漏电流 J_ℓ (nA cm^{-2})从集电极流向基极，如图6.4所示。

第5章曾经提到反偏pn结电流的三种成分：空间电荷区的产生流，p型基区中的电子产生流和n型集电区中的空穴产生流。从等式(5.3.29)中可以看出，空间电荷区的产生流是硅pn结反向电流的主要成分。

由5.3节知，空间电荷区产生电流 I_g 大约与空间电荷区的宽度成正比，与集电极-基极偏压的依赖关系相对较弱(等式(5.3.26)～(5.3.28))，通常可以写成

$$I_\ell = J_\ell A_C = I_g = \frac{1}{2}\frac{qn_i x_d A_C}{\tau_0} \tag{6.3.1}$$

式中 x_d 是空间电荷区宽度、A_C 是集电结面积，τ_0 为空间电荷区的有效寿命。由于 I_ℓ 的值很小，因此在硅晶体管工作于放大区时可以忽略，但是在晶体管开关工作中很重要。根据图6.3所示的集成电路晶体管的结构，面积 A_C 约等于 $Y_C \times Z_C$ 加上靠近表面的集电结的纵向部分。表面复合(5.3节)也是 I_ℓ 的一个组成部分，经过严格计算可发现这个组分非常小。

本节主要讨论了图6.7中的晶体管的各工作区。晶体管开关交替工作在区域1和区域3(饱和与截止),区域2和区域4为开关过程的瞬时状态。工作状态从饱和到截止时,存储在基区及其附近的电荷发生改变。开关电路的设计要求保证充分的充电和放电,第7章将继续讨论符合晶体管物理过程并适合于计算的模型。下一节将介绍适用于直流应用的晶体管模型并量化前面提到的物理图形。

6.4 EBERS-MOLL 模型

J. J. Ebers 和J. L. Moll 于1954年提出了一个简单有用的模型来表示双极晶体管中载流子的注入和抽取。四十多年后,这个模型为双极晶体管(BJT)复杂的计算机辅助计算提供了基本的框架。6.1节用两个背靠背二极管来理解双极晶体管就是以Ebers-Moll模型(EM模型)为基础的,并且由此推导出了等式(6.1.14)。对于放大偏置晶体管,如图6.5所示,等式(6.1.14)中的 J_n 代表从发射结流向集电结的电流,它把发射区和集电区联系到一起,因此 J_n(相应的pnp晶体管中的 J_p)也被称做连接电流。6.1节介绍的理论中只考虑了连接电流,而没有考虑图6.5中的基极电流组分。

研究基极电流时,可分别考虑基极与发射极之间的电流 I_{BE},以及基极与集电极之间的电流 I_{BC}。由于发射结是一个pn结二极管,因此从基区流向发射区的空穴电流可以用理想二极管方程来表示,饱和电流用 I_{0E} 表示

$$I_{BE} = I_{0E}[\exp(qV_{BE}/kT) - 1] \tag{6.4.1}$$

发射极总电流等于流向集电极的电流(连接电流)减去 I_{BE}。由等式(6.1.14)可求出连接电流。假设基区横截面积为常数 A,于是有 $I_S = J_S A$,这里的 J_S 可以由等式(6.1.15)给出[9]。因此,发射极电流为

$$I_E = I_S[\exp(qV_{BC}/kT) - \exp(qV_{BE}/kT)] - I_{0E}[\exp(qV_{BE}/kT) - 1] \tag{6.4.2}$$

同样,可写出集电极电流

$$I_C = I_S[\exp(qV_{BE}/kT) - \exp(qV_{BC}/kT)] - I_{0C}[\exp(qV_{BC}/kT) - 1] \tag{6.4.3}$$

其中基极与集电极间的电流为

$$I_{BC} = I_{0C}[\exp(qV_{BC}/kT) - 1] \tag{6.4.4}$$

根据对电压的依赖关系,可以将等式(6.4.2)和(6.4.3)重组为

$$I_E = -(I_S + I_{0E})[\exp(qV_{BE}/kT) - 1] + I_S[\exp(qV_{BC}/kT) - 1] \tag{6.4.5a}$$

和

$$I_C = -(I_S + I_{0C})[\exp(qV_{BC}/kT) - 1] + I_S[\exp(qV_{BE}/kT) - 1] \tag{6.4.5b}$$

其中定义

$$I_{ES} \equiv I_S + I_{0E}, \qquad I_{CS} \equiv I_S + I_{0C} \tag{6.4.6a}$$

和

$$\alpha_F \equiv \frac{I_S}{I_S + I_{0E}}, \qquad \alpha_R \equiv \frac{I_S}{I_S + I_{0C}} \tag{6.4.6b}$$

[9] 在Ebers-Moll模型中基区面积A不需要为常数(非集成电路晶体管)。

将等式(6.4.6)代入等式(6.4.5)可以得到

$$I_E = -I_{ES}[\exp(qV_{BE}/kT) - 1] + \alpha_R I_{CS}[\exp(qV_{BC}/kT) - 1] \qquad (6.4.7a)$$

和

$$I_C = -I_{CS}[\exp(qV_{BC}/kT) - 1] + \alpha_F I_{ES}[\exp(qV_{BE}/kT) - 1] \qquad (6.4.7b)$$

等式(6.4.7)为 npn 晶体管的 Ebers-Moll(EM)方程。相应的 pnp 晶体管的 EM 方程中,电流方向有所改变以说明 pn 结极性的改变。当 V_{EB}与 V_{CB}为正时,pnp 双极晶体管中的两个二极管处于正偏状态(习题 6.10)。

由 EM 方程可以直接得出晶体管的发射极电流和集电极电流,结合基尔霍夫电流定律(流经同一点的所有电流之和为零),还可以求出基极电流。EM 模型有四个参数(α_F,α_R,I_{ES}和I_{CS}),由等式(6.4.6)知只有三个独立的参数,第四个参数可根据互易关系得出。

$$\alpha_F I_{ES} = \alpha_R I_{CS} = I_S \qquad (6.4.8)$$

引入两个新的参量:正向放大偏置下的二极管电流 I_F和反向放大偏置下的二极管电流 I_R

$$I_F = I_{ES}[\exp(qV_{BE}/kT) - 1] \qquad (6.4.9a)$$

和

$$I_R = I_{CS}[\exp(qV_{BC}/kT) - 1] \qquad (6.4.9b)$$

等式(6.4.7)可以简化为

$$I_E = -I_F + \alpha_R I_R \qquad (6.4.10a)$$

和

$$I_C = -I_R + \alpha_F I_F \qquad (6.4.10b)$$

与等式(6.4.10)对应的等效电路为图 6.12。电路包括基极和发射极、基极和集电极之间的二极管和电流源,电路中的电流源用于表示与较远结上的偏压有关的电流(发射结处的 I_R和集电结处的 I_F)。对图 6.12 中的电路使用基尔霍夫电流定律,可以求出基极电流为

$$I_B = -(I_E + I_C) = I_F(1 - \alpha_F) + I_R(1 - \alpha_R) \qquad (6.4.11)$$

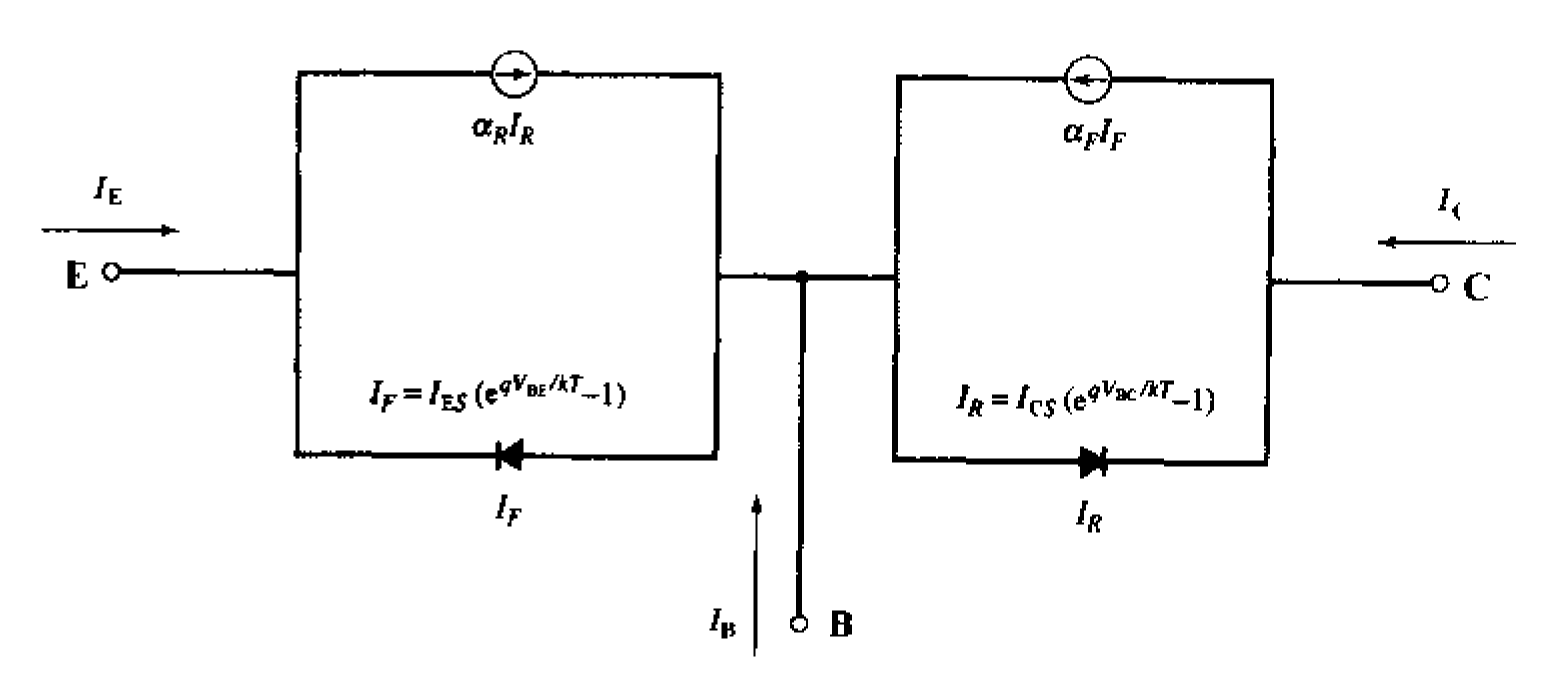


图 6.12　npn 晶体管 EM 模型的等效电路

应用 下面介绍怎样用EM模型表示晶体管不同的工作区。首先考虑 V_{BE} 与 V_{BC} 都为负的截止区。由等式(6.4.9)、(6.4.10)和(6.4.11)可以得到如图6.13所示的等效电路,模型被简化为只有两个表示结反向饱和电流的电流源。

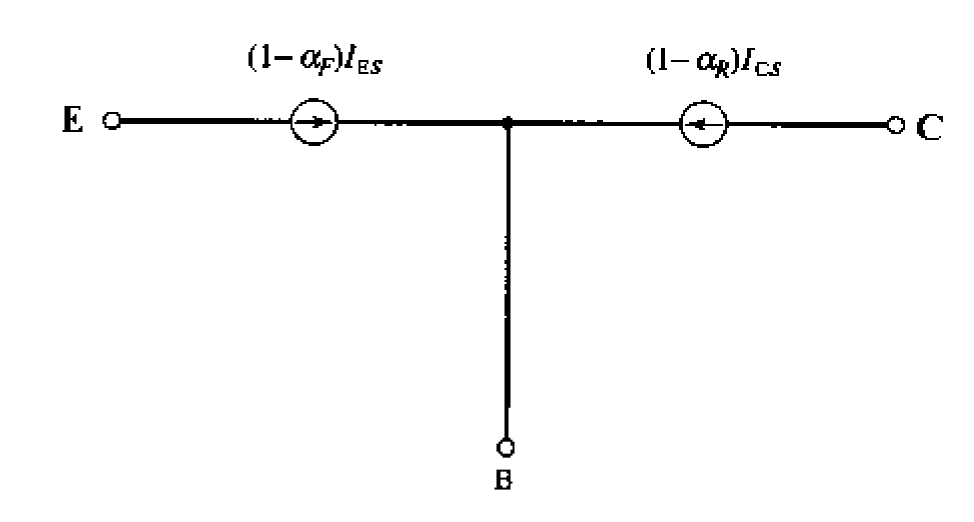


图6.13 晶体管截止时的EM模型

在正向放大区,发射结正偏而集电结反偏,可以用发射极电流来表示集电极电流,则等式(6.4.10)可以被写成

$$I_C = -\alpha_F I_E - I_R(1 - \alpha_F\alpha_R) \tag{6.4.12}$$

在放大偏置条件下为

$$I_C = -\alpha_F I_E + I_{CS}(1 - \alpha_F\alpha_R) \tag{6.4.13}$$

同样,在反向放大偏置条件下有

$$I_E = -\alpha_R I_C + I_{ES}(1 - \alpha_F\alpha_R) \tag{6.4.14}$$

从等式(6.4.13)与(6.4.14)发现,EM模型参数 α_F、α_R、I_{ES} 和 I_{CS} 可以通过测量正向放大偏置下 I_C 与 I_E 的关系,或者反向放大偏置下 I_E 与 I_C 的关系得到。在正向放大偏置下,I_C 与 I_E 呈线性关系,且斜率等于 $-\alpha_F$。$I_E=0$ 处的截距对应发射极处于开路的情况,这时的电流通常用 I_{CB0} 表示

$$I_{CB0} = I_C|_{I_E=0} = I_{CS}(1 - \alpha_F\alpha_R) \tag{6.4.15}$$

(下标的前两位表示电流流经的两个电极,第三位表示第三个极(此时是发射极)开路)。

这个电流可与基极开路时的集电极电流 I_{CE0} 对比。$I_E = -I_C$ 时,由等式(6.4.13)得出

$$I_{CE0} = I_C|_{I_B=0} = \frac{I_{CS}(1 - \alpha_F\alpha_R)}{(1 - \alpha_F)} = \frac{I_{CB0}}{(1 - \alpha_F)} \tag{6.4.16}$$

I_{CB0} 与 I_{CE0} 的不同可以追溯到发射结的偏压状态。当测量 I_{CB0} 时,由于发射区的部分电子被抽走,又没有得到外部电流的补偿,因此发射结轻微反偏,此时集电极电流中只有基区产生的电子和集电区产生的空穴。测量 I_{CE0} 时,基极开路,发射结正偏,集电极大部分电流可由从发射结注入到基区的电子组成,因此晶体管增益将有效地提高漏电流 I_{CE0}。

例题 Ebers – Moll 方程

已知npn双极晶体管,发射极开路,集电结反偏,计算发射结反偏电压。假设 $\alpha_F = 0.98$,$\alpha_R = 0.70$,$I_{CS} = 1\times10^{-13}\,\text{A}$,$I_{ES} = 7.14\times10^{-14}\,\text{A}$。

解:在这种偏压条件下,集电极电流等于等式(6.4.15)给出的 I_{CB0}。因为发射极电流为零,所以等式(6.4.10(a))可以写成

$$I_F = \alpha_R I_R \approx -\alpha_R I_{CS} = I_{ES}\left(\exp\frac{qV_{BE}}{kT} - 1\right)$$

其中利用了 $I_R \approx -I_{CS}$。

利用互易关系(等式(6.4.8))可以解出 $T = 300\text{K}$ 时的 V_{BE}

$$V_{BE} = \frac{kT}{q}\ln(1 - \alpha_F) = -0.10\ \text{V}$$

在发射极开路的情况下,发射结反向偏压只与 α_F 有关,这是因为该偏压是为了平衡连接电流与反偏发射结二极管电流,使发射极电流为零而产生的。

利用端电流驱动的有源元件来获得模型的方法非常有用。放大偏置下利用以发射极电流为变量的方程很容易得到这种模型。图6.14(a)是与等式(6.4.13)相符的发射极驱动模型的

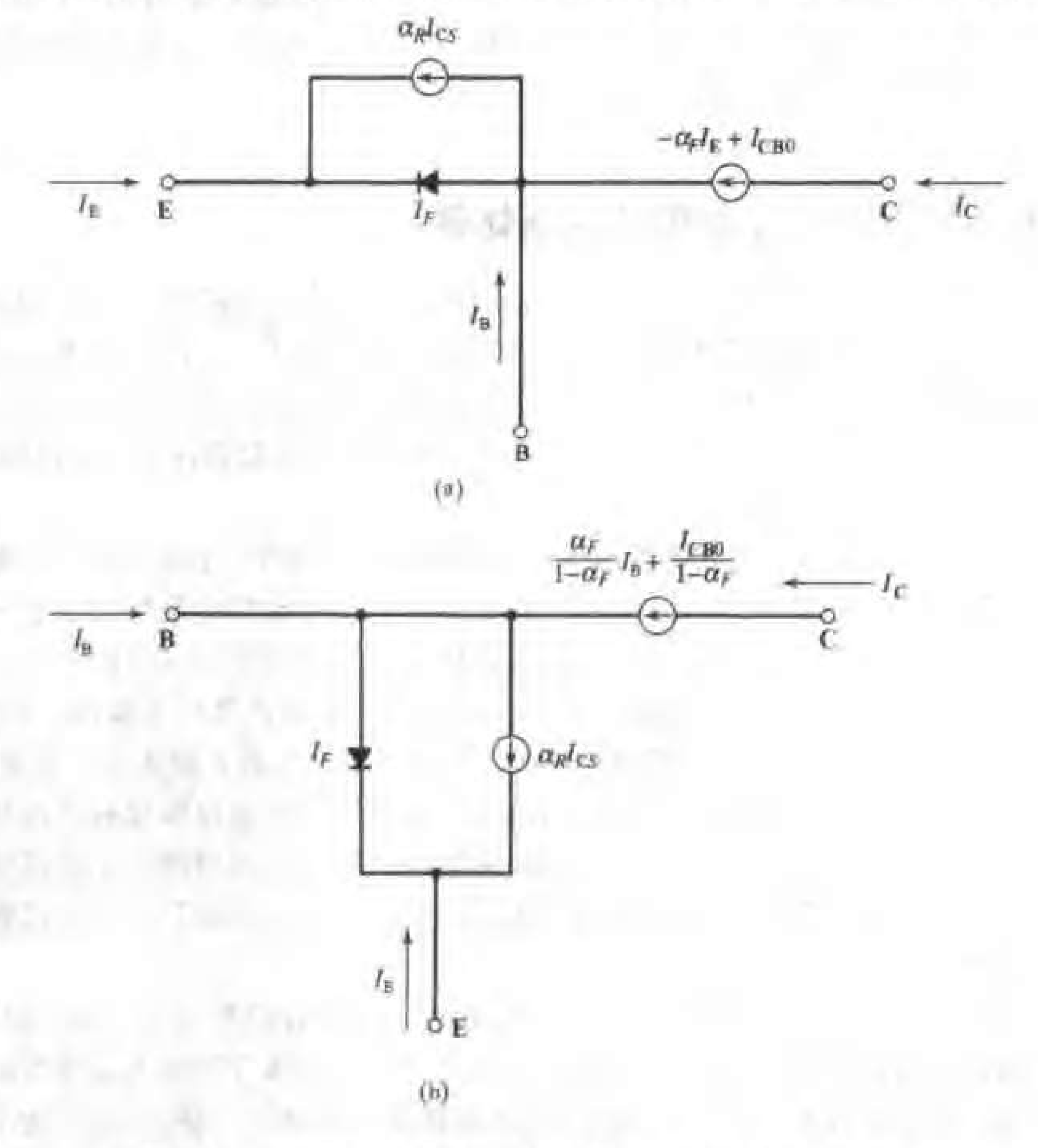


图6.14　由端电流驱动的npn晶体管等效电路。(a)发射极电流驱动等效电路;(b)基极电流驱动等效电路

等效电路,其中集电结漏电流用 I_{CB0} 表示(等式(6.4.15))。也可以直接推导出放大偏置晶体管的基极电流驱动的表达式(习题6.11),并证明图6.14(b)的电路。

饱和区最受关注的量是"开态"晶体管上的压降 V_{CEsat}。由 EM 模型可得出

$$V_{CEsat} = \frac{kT}{q}\ln\left\{\frac{\left[1+\frac{I_C}{I_B}(1-\alpha_R)\right]}{\alpha_R\left[1-\frac{I_C}{I_B}\left(\frac{1-\alpha_F}{\alpha_F}\right)\right]}\right\} \tag{6.4.17}$$

推导等式(6.4.17)时,考虑到饱和晶体管的二极管电流表达式中的指数部分远大于1(习题6.12)。因为 V_{CEsat} 很小,一级近似通常可忽略。不能被忽略时,等式(6.4.17)指出 V_{CEsat} 只随集电极电流缓慢变化,因此,等效电路中可以用从集电极到发射极的电压源等效。更可取的方法是用从基极到发射极和从基极到集电极的两个电压源等效。EM 方程没有考虑与结串联的电阻的影响。串联电阻上的压降,尤其是集成电路晶体管的集电区串联电阻,经常使 V_{CEsat} 超过等式(6.4.17)预期的值,因此,"开态"晶体管模型通常用电压源和电阻 R_{Csat} 串联组成。

本节讨论了直流条件下的静态 EM 模型。虽然这个模型经修正也可用于动态计算(即解决瞬态问题),但电荷控制模型更适合于这类计算。电荷控制模型将在第7章介绍,同时也会对 EM 模型进行修改,使其能够解决一些重要的二级效应。

6.5 器件:平面双极放大和开关晶体管

在讨论集成电路双极晶体管时,可以根据它们的基本用途分为两大类:放大晶体管和开关晶体管。由于电子的迁移率和扩散系数高于空穴,npn 晶体管比 pnp 晶体管更为常用。放大和开关 npn 晶体管经常被制作在电阻率相对较高的 n 型外延层上。为了防止在集电结工作电压范围内的电击穿(4.4节),外延层一般为轻掺杂。外延层通常淀积在 p 型硅衬底上,这样晶体管就能通过 pn 结与衬底隔离。

如图6.3(b)所示,集成电路晶体管的集电极接触位于晶体管上表面的放大区附近。为了减小由金属电极流入的横向电流上的串联电阻,需要在轻掺杂 n 型外延层与 p 型衬底之间加入一个重掺杂 n 型埋层或次集电极,这个问题我们已经讨论过,如图5.18(b)所示。一般晶体管中的埋层可以将串联电阻从几千欧姆减小至几百欧姆。相邻的器件在横向方向相互隔离,这将在下一节讨论。基区、发射区和集电极接触区依次通过注入或扩散形成。需要指出的是,放大晶体管即使有埋层也不能足够地减小串联电阻,因为埋层和集电极接触之间的外延层电阻仍然很大。这时可加入一个额外的工艺步骤,在集电极接触下面增加一个重掺杂 n 型区且一直扩散至掩埋层。这个延长的扩散区被称做集电极插塞(Collector Plug),可以将串联电阻减小至10Ω 量级。

图6.15与图6.16分别为典型的开关晶体管和放大晶体管的垂直于表面方向,经过发射区、基区和集电区的杂质浓度分布图。从图中可知设计开关晶体管和放大晶体管的主要区别在于外延层的厚度与电阻率。这两个参数在放大晶体管中都要大一些,目的是为了增大击穿电压和减小寄生效应(Early 效应),这些将在第7章进一步讨论。开关器件要求尽量减小饱和

（“开态”）电阻，这就要求外延层厚度大约为一微米左右，而电阻率大约在零点几 Ω·cm 左右[10]。如果开关晶体管中考虑 Schottky 箝位（见第 3 章），电阻率必须大于 0.1Ω·cm 以获得较好的金属-半导体势垒。Schottky 箝位晶体管与常规非箝位晶体管惟一的区别在于，它的基极电极接触区延伸到了轻掺杂外延集电区，如图 6.17 所示。为了增大发射效率 γ，需要减小基区掺杂浓度。但是如果基区掺杂浓度太低，基区横向串联电阻会限制晶体管的性能。如果为获得合适的发射效率而使晶体管有源基区（本征基区）掺杂足够低，则需要在有源基区以外（非本征基区）的基区掺杂 p 型材料来减小横向串联电阻和提高金属-半导体接触性能。这种做法还可以避免有源基区以外的 p 型区由于轻掺杂而发生表面反型（见第 8 章）。

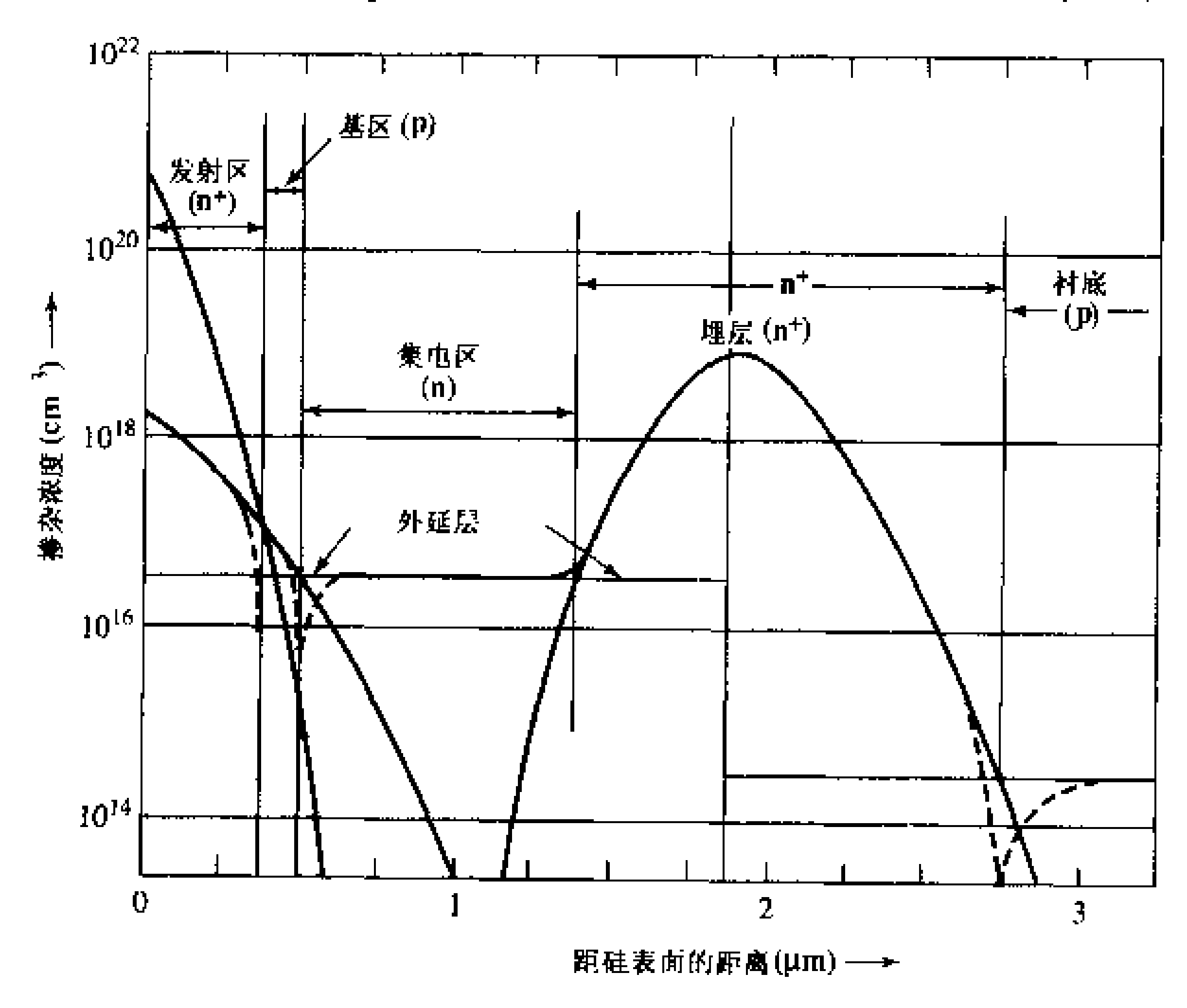


图 6.15　开关晶体管的掺杂分布[8]

发射区通常重掺杂以提高发射效率，但当掺杂浓度高于 $10^{20}\mathrm{cm}^{-3}$ 时，发射区空穴寿命减小（这将增加从基区注入的空穴）以及禁带变窄（1.1 节）将导致发射效率下降[16]。

前面已经提到，减小基区宽度和降低基区掺杂浓度都能提高增益。然而当基区宽度减小至低于微米量级时，外加的基极-集电极电压足够耗尽整个中等掺杂的中性基区，如图 6.18 所示，导致集电结空间电荷区和发射结空间电荷区相连。这时由于发射结势垒降低，发射极到集电极的电流急剧增加，出现穿通，将破坏晶体管正常的工作状态。

另一个非正常状态是 V_{CB} 引起的集电结雪崩击穿，即 4.4 节介绍的二极管击穿，但是双极晶体管的击穿电压可能远低于反偏的集电极-基极二极管。考虑基极开路的 npn 双极晶体管，发射极到集电极上的压降使得集电结反偏，大部分外加电压降落在集电结上。为了提供集电结的反向饱和电流，需要从集电区向基区注入少量的空穴。空穴作为基区的多数载流子，无法

[10] 双极晶体管饱和时，集电极与发射极之间的电压降等于等式（6.4.17）给出 $V_{CE,sat}$ 加上由于晶体管结构形成的串联电阻上的电压降。通常最大的串联电阻位于外延层。

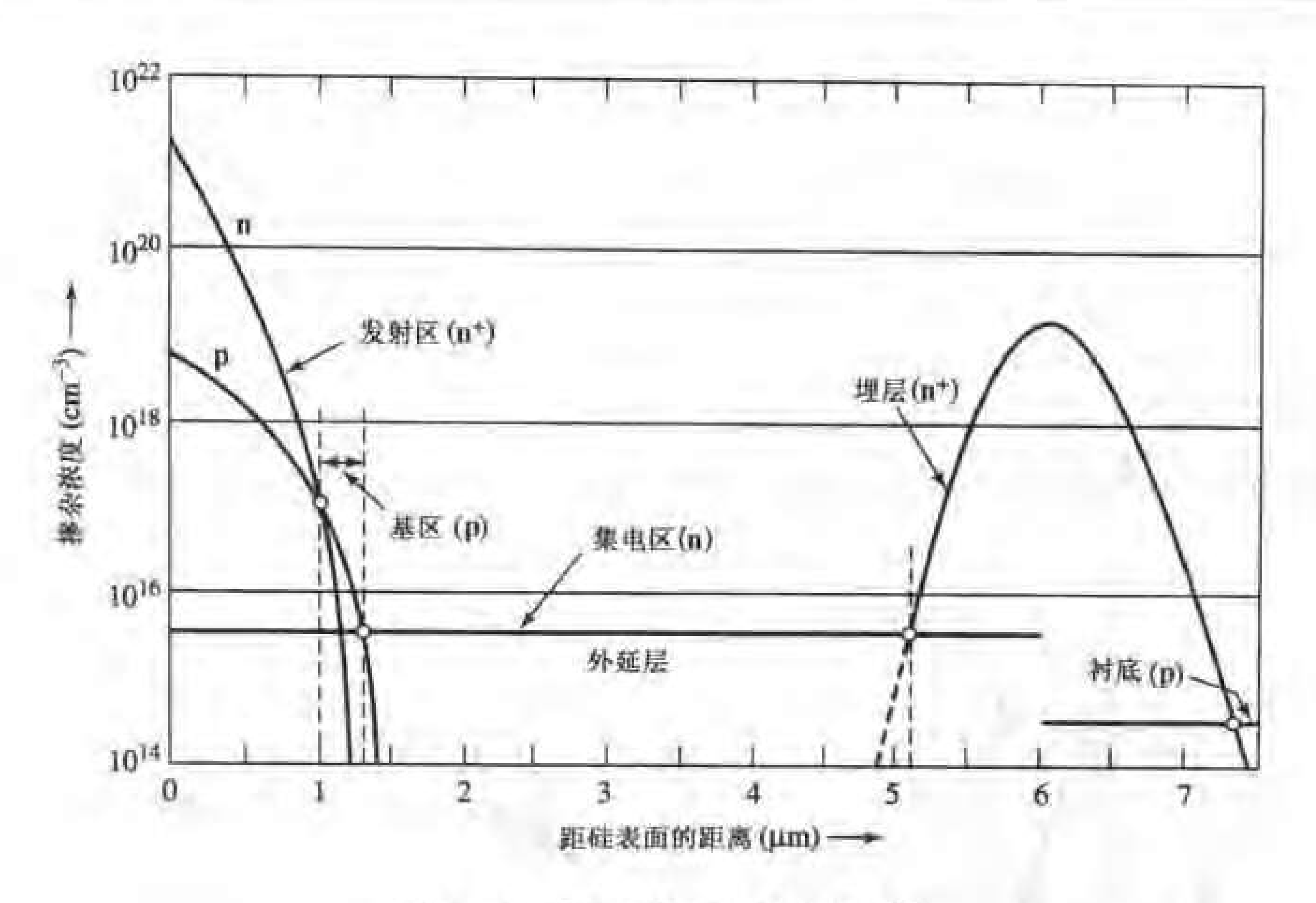


图 6.16 放大晶体管的掺杂分布[8]

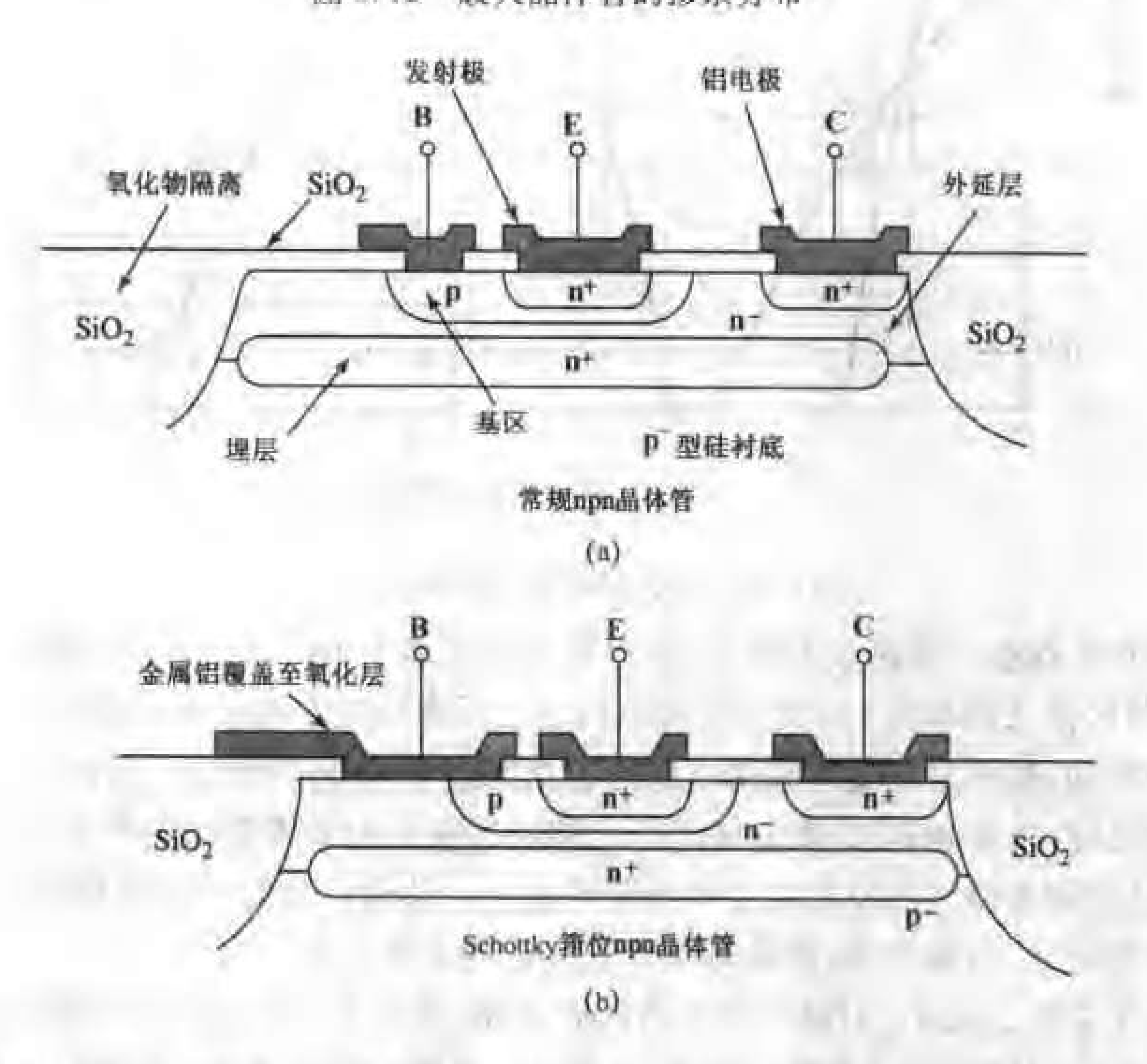


图 6.17 (a)常规 npn 晶体管和(b)Schottky 箝位 npn 晶体管的横截面示意图

在中性基区积累而移至中性基区边界,中和了一部分发射结空间电荷区,使得发射结空间电荷区减薄,势垒降低,导致电子从发射区注入到基区,并穿过基区到达集电结空间电荷区。反偏集电结的强电场加速了进入集电结空间电荷区的电子,而这些电子可以离化出其他电子,最终导致集电结雪崩击穿。

发射结势垒高度的微小下降就能导致从发射结注入(和运动到集电结)的电子数目急剧

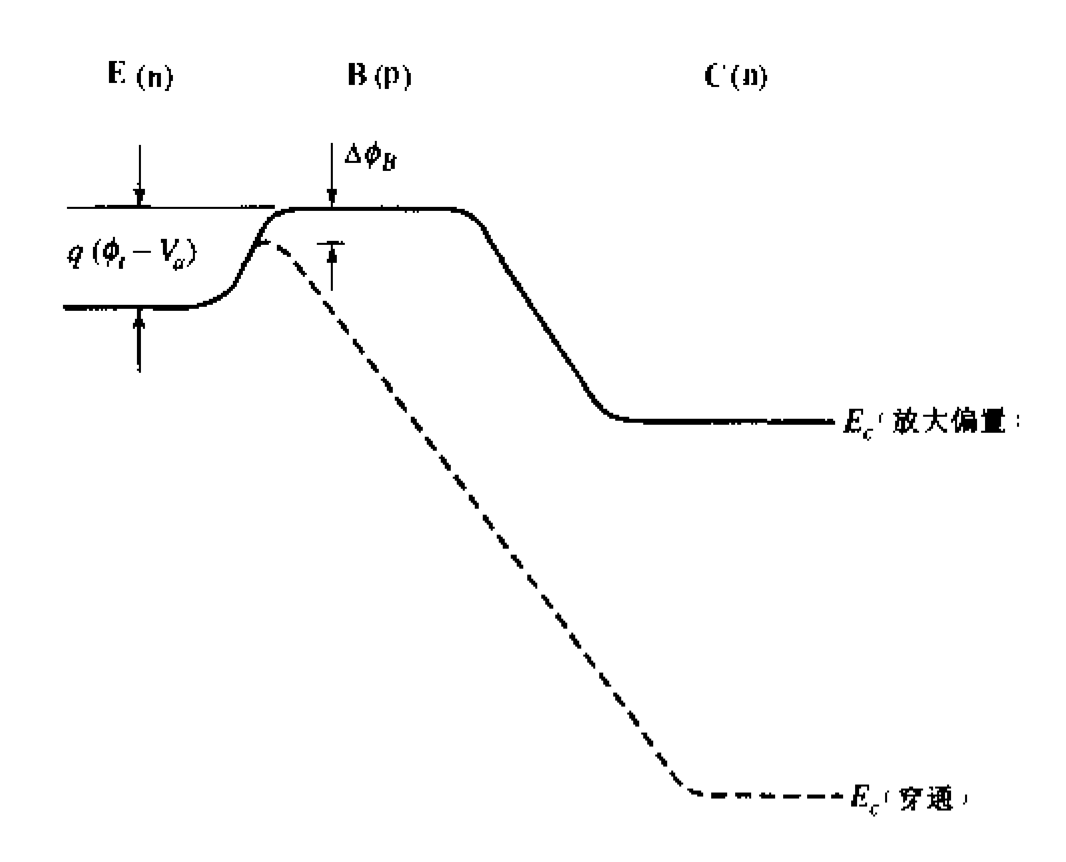


图6.18　实线:npn晶体管正常放大偏置下的导带边能带图;虚线:集电结偏压很高,整个基区中的可动载流子全部耗尽,导致基区穿通,并导致发射结电子势垒下降了$\Delta\phi_B$

升高,并且进入集电结的电子数目远大于基区平衡态电子(少数载流子)(电子数目被晶体管有效地放大了β倍)。由于有大量电子,基极开路晶体管的集电结雪崩击穿电压远低于基极-集电极二极管的击穿电压。这个击穿电压BV_{CEO}(基极开路时的击穿电压)与基极-集电极二极管击穿电压(BV_{CBO})之间的关系为

$$BV_{CEO} = \frac{BV_{CBO}}{\beta^{1/m}} \tag{6.5.1}$$

式中m的典型值约为4。

工艺考虑

本征晶体管的性能主要决定于纵向掺杂分布,但本征晶体管附近的寄生元件也会严重影响晶体管特性。非本征电阻和电容影响了晶体管的增益和频率响应特性。

集成电路中晶体管之间以及晶体管与衬底之间必须相互隔离。晶体管与衬底之间的纵向隔离通过n型外延集电区和p型衬底之间的pn结实现。早期的集成电路工艺中,通常在相邻晶体管之间的n型外延层插入一个p型区来实现横向隔离,如图6.19(a)和6.20(a)所示。这个扩散隔离区必须穿透整个外延层延伸至p型衬底。由于横向扩散的影响,隔离区的横向尺寸通常较大,限制了集成度。另外,p型基区必须远离p型隔离区,这样才能使它们能被中性n型区分开,并且从基区(横向寄生双极晶体管的发射区,该寄生极晶体管由p型基区、n型集电区和p型隔离区构成)注入的载流子不会被隔离区收集。

使用氧化物横向隔离可以大幅度减小晶体管的横向尺寸。如果外延层的厚度为1μm量级,那么用2.6节介绍的LOCOS(硅的局部氧化)工艺就可以实现氧化层穿透整个外延层。

对于稍厚一些的外延层,首先要将被氧化区域的硅外延层刻蚀掉大约一半厚度,这样生长

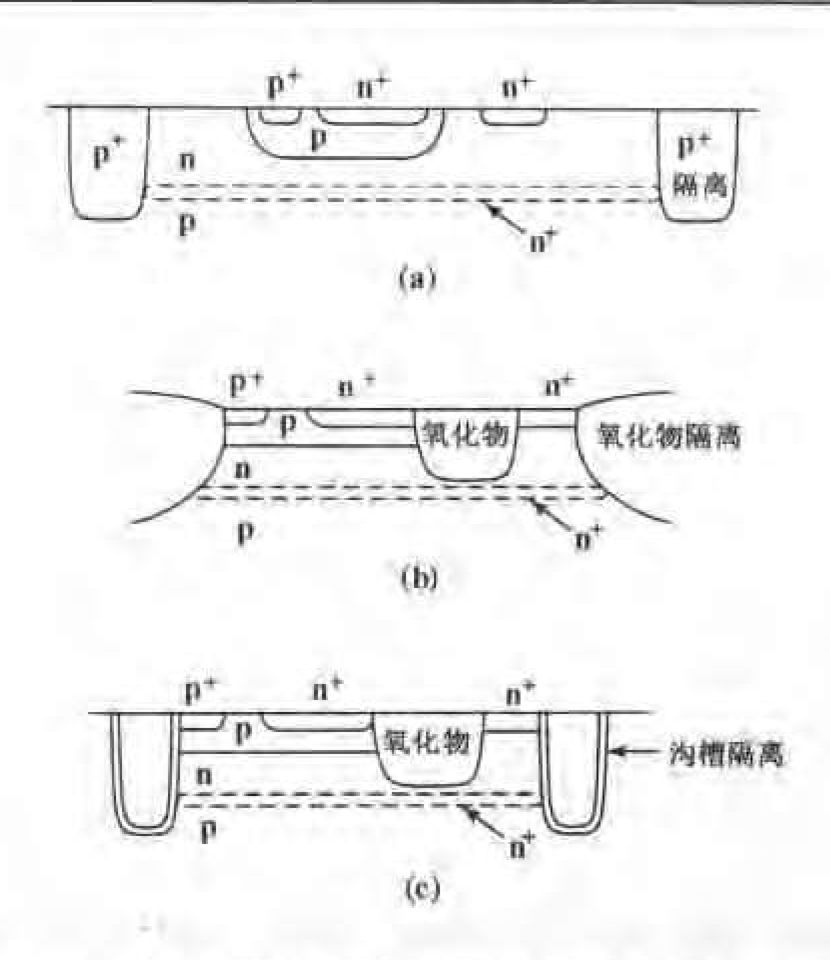


图6.19 包括隔离区的BJT横截面示意图 (a) pn结隔离；(b) LOCOS氧化物隔离；(c)沟槽隔离

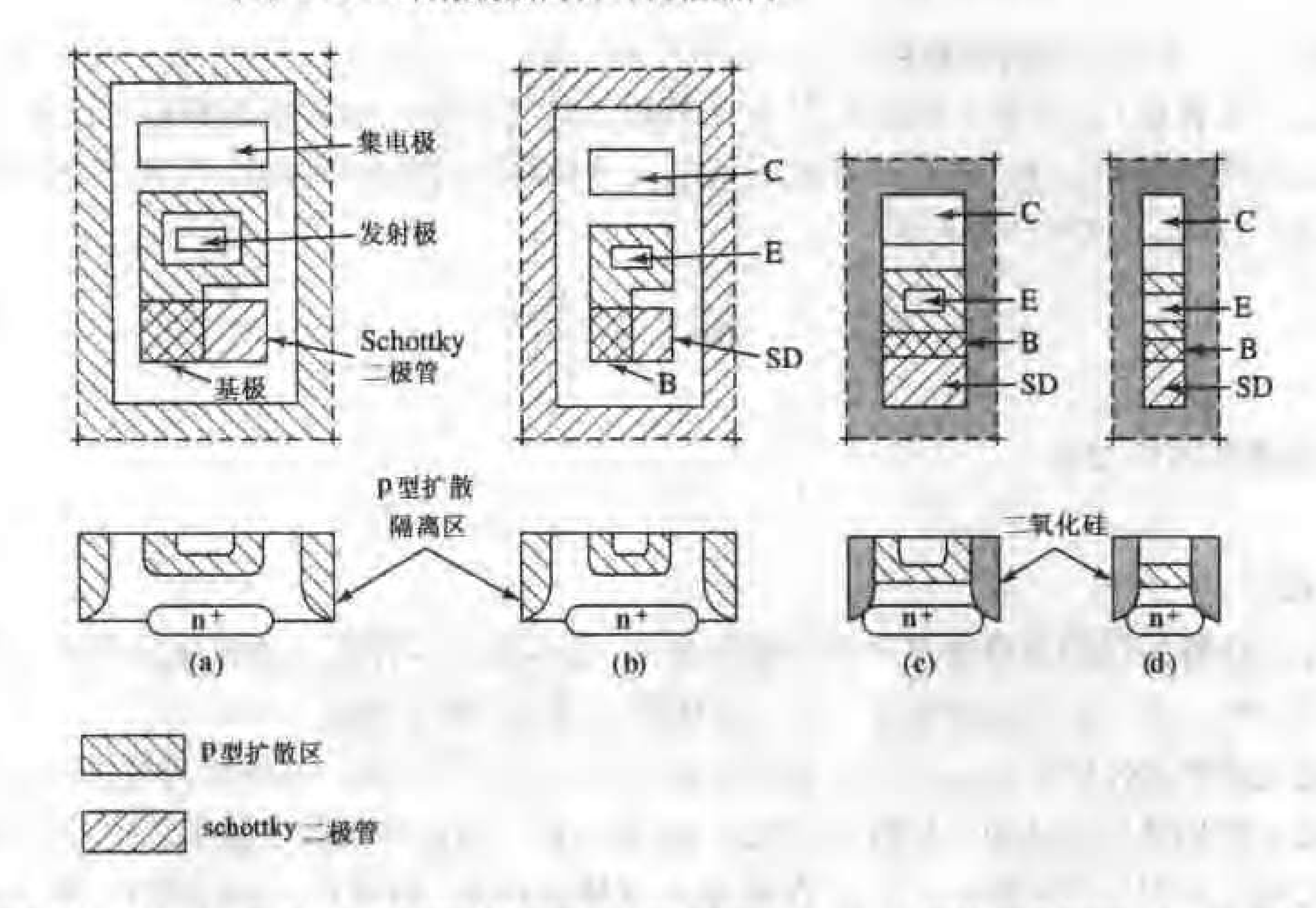


图6.20 Schottky箝位双极晶体管的俯视图与横截面图。(a) pn结隔离晶体管；(b) pn结隔离泡发射区(washed-emitter)晶体管；(c) 氧化物隔离晶体管；(d) 墙围发射区(walled-emitter)晶体管。图中用阴影标出了p型扩散区与Schottky二极管，并指出了各电极接触区

氧化物的区域将低于表面，如图6.19(b)所示。当在表面生长氧化层时，隔离区的氧化物厚度与整个外延层厚度一致，而不是一半(生长氧化物时，氧化物一半在原始表面上方，一半在原始表面下方)。

对于更厚的外延层，可以使用沟槽隔离，如图6.19(c)所示。用定向离子刻蚀法在整个外

延层纵向刻出一个"沟槽"(过刻蚀一小段厚度以补偿工艺偏差),接着在沟槽的侧面和底部生长氧化层。沟槽的其余部分填充未掺杂多晶硅以保持表面平整(如果外延层很薄,沟槽中可以完全填充绝缘氧化物,但深沟槽如果全部填充氧化物会产生应力,导致晶片变形,使后续工艺变得复杂)。沟槽隔离工艺不仅使用灵活,还能增加集成度,因而避免了LOCOS工艺造成的器件与隔离区之间的横向氧化。

为缩小晶体管的面积,可以省去制作发射极接触孔的掩膜版(同时也省去了与对准工艺相关的额外面积),通过直接刻蚀被薄氧化层覆盖的发射区来形成电极接触区,然后蒸发金属覆盖整个发射区;接触面积的增加还可减小寄生电阻。这种泡发射极工艺如图6.20(b)所示,要求发射区以外的其他区域氧化层较厚,不会被刻穿。由于发射区的掺杂杂质在覆盖基区的厚氧化层下方横向扩散,所以(大部分时候)可以避免基极-发射极短路。

当用氧化物作隔离区时,基区和发射区不需要远离隔离区,可紧挨着隔离区,如图6.20(d)所示,这样可进一步减小晶体管的面积。由于二氧化硅的介电常数低于硅,减小了寄生侧壁电容。为进一步节省面积,还可以用一个较窄的氧化层隔离发射区和集电极电极接触区。

多晶硅发射极工艺 大部分先进的双极晶体管采用了6.2节提到的多晶硅发射极接触。图6.21为双层多晶硅工艺流程:首先形成埋层、外延层和LOCOS或沟槽隔离,接着淀积第一层多晶硅,并重掺杂p型杂质。这层多晶硅的一端与单晶硅基区有很少的接触,另一端将在厚氧化层上方与金属电极接触,较低的二氧化硅介电常数减小了基极接触区的寄生电容。部分p型杂质从多晶硅扩散到单晶硅的非本征基区接触区,在多晶硅和单晶硅之间形成欧姆接触,多晶硅的上方和侧面形成了氧化层。去除本征基区上方的氧化层,通过注入形成本征基区。

接下来淀积第二层多晶硅,通过氧化层与第一层多晶硅隔离。自对准工艺减小了非本征基区的寄生电阻。如果前面没有对本征基区掺杂,那么可将杂质注入到第二层多晶硅,并从多晶硅扩散到单晶硅的基区中。然后向多晶硅注入发射区杂质,使其扩散到单晶硅中形成一层很薄的发射区。不直接向单晶硅注入可以避免晶格损伤和由此增强的杂质扩散。厚度大约为50~100nm量级的基区和发射区可以采用这种结构。发射极和集电极接触可放在氧化层上方n型多晶硅的延伸部分,正如基极接触位于氧化层上方的p型多晶硅延伸部分。这样的结构可以减小寄生电容,优化器件的交流性能。

总则 图6.22(a)和6.22(b)为双极晶体管的符号,其中两条斜线代表集电极和发射极,与之相交的直线代表基极。发射极-基极二极管从p区指向n区的箭头用以区分npn晶体管(见图6.22(a))与pnp晶体管(见图6.22(b))。实际上,集成电路中的npn晶体管通常与一个以衬底为集电极的寄生pnp晶体管耦合(见图6.22(c))。当npn晶体管正向放大偏置时,衬底pnp晶体管处于截止状态。npn晶体管处于其他工作模式时,pnp晶体管可能进入放大工作状态。为避免寄生损耗和误操作,必须保证寄生pnp晶体管几乎没有增益,此时埋层的作用非常重要。埋层通过增加寄生pnp晶体管的基区电荷,减小了基区载流子寿命,降低了pnp晶体管的α_F(一般情况下低于0.05)。对于pn结隔离双极晶体管,寄生pnp晶体管效应还包括侧向注入的空穴经过外延层,被反向偏置的p型隔离区收集。实际上,保证基区外边界与隔离区分离是非常必要的,这样可以将该寄生元件的影响减小到可以接受的程度。如果使用氧化物隔离,则没有这个问题。

从发射区到基区也存在着电子的横向注入,我们在讨论晶体管增益时,没有考虑这部分注

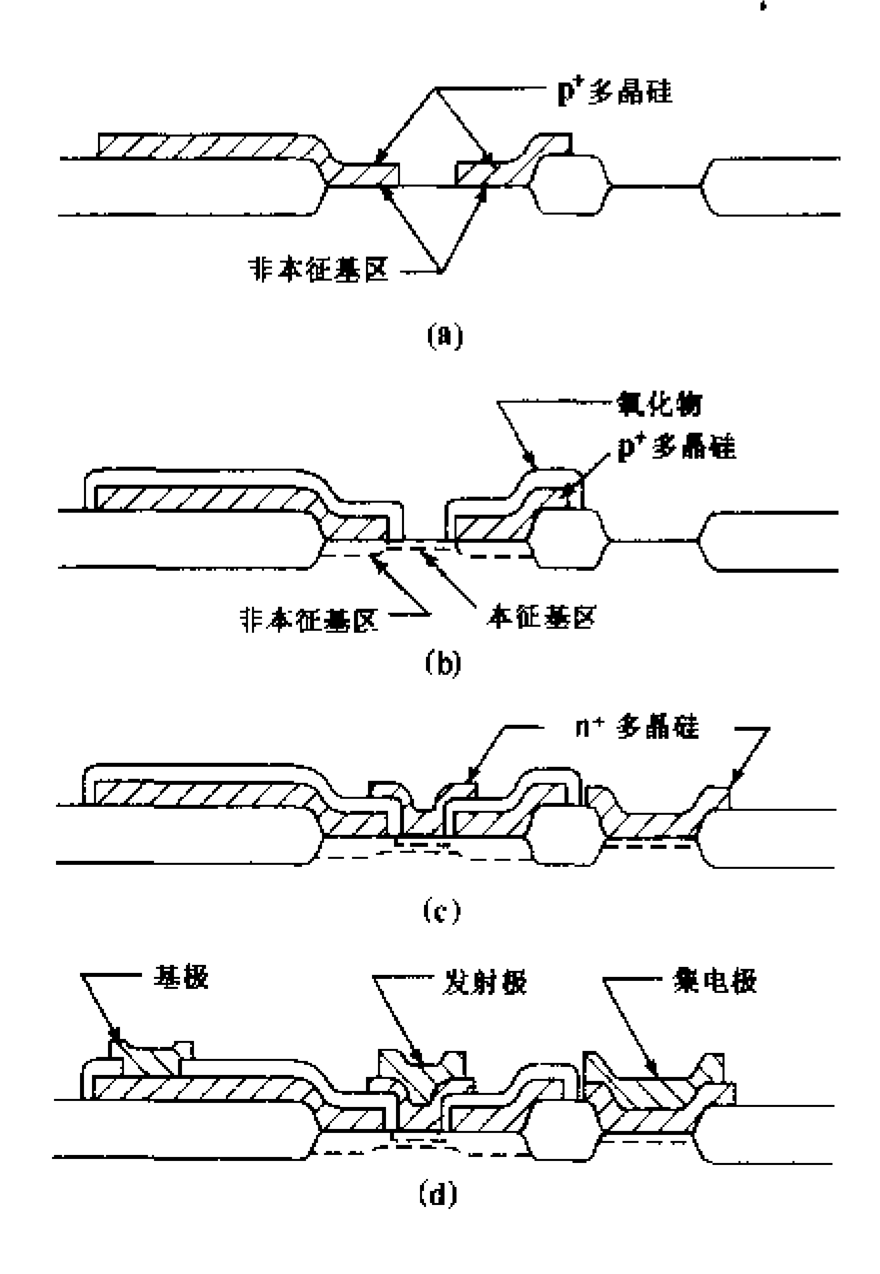


图6.21 双层多晶硅晶体管工艺流程,两层多晶硅之间用氧化物隔离。(a)第一层多晶硅形成非本征基区,并延伸至场氧化层;(b) 向多晶硅注入硼杂质,通过硼杂质向单晶硅的扩散形成与本征基区的接触。扩散的同时多晶硅和单晶硅被氧化。去除单晶硅上方的氧化层可用于形成发射区;(c)淀积第二层多晶硅并掺n型杂质,通过热处理形成最后的发射区和基区杂质分布;(d)在氧化层上方形成基极接触,可以减小单晶硅中的非本征基区面积[14]

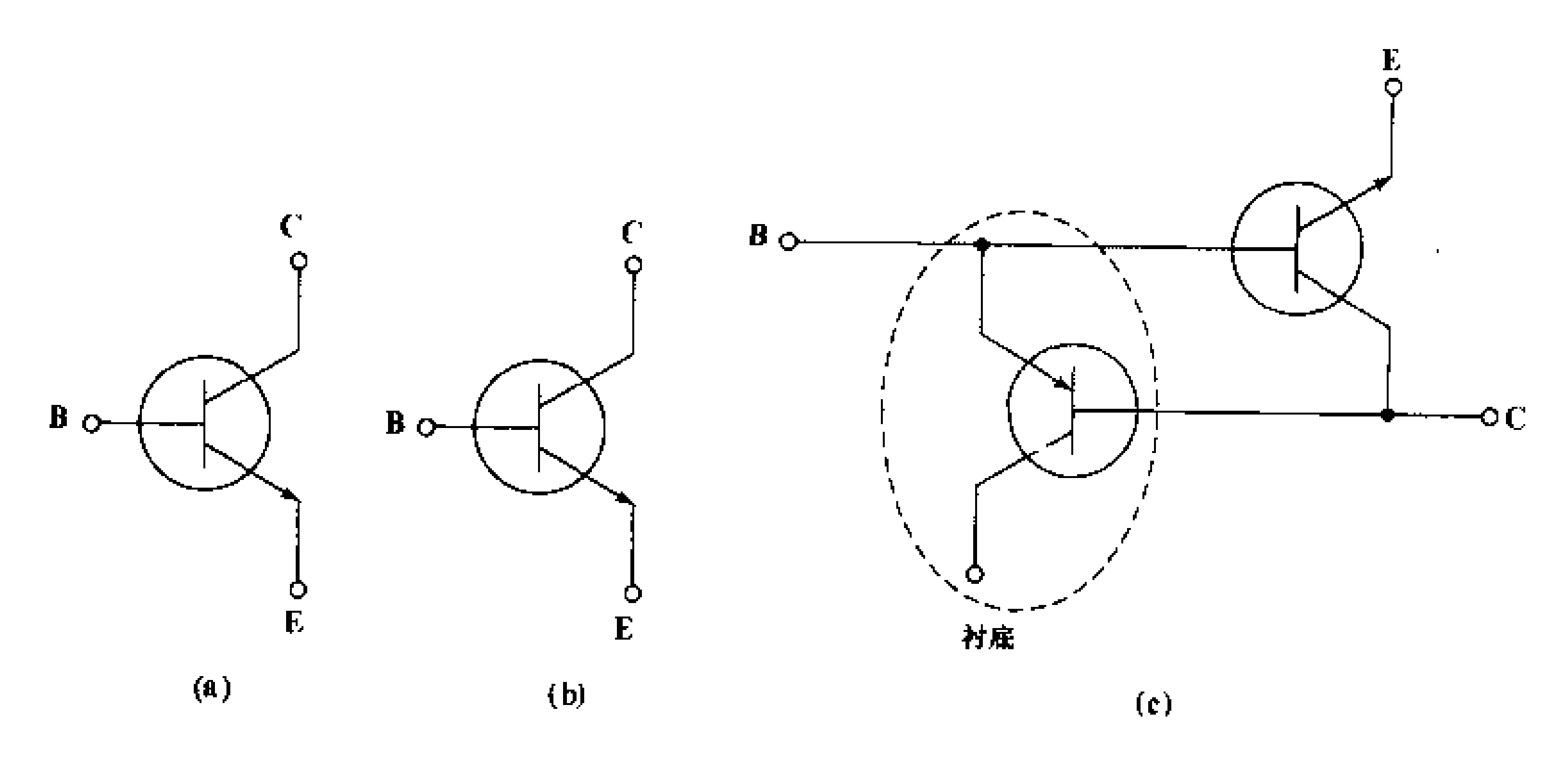


图6.22 (a)npn晶体管和(b)pnp晶体管的标准符号;(c)由制作工艺引起的包含寄生pnp晶体管(虚线框出)的IC npn晶体管

入造成的增益损失。但是随着集成电路晶体管尺寸的减小,横向注入损失不可忽略。由于有效的横向基区宽度很大,并且部分电子在基极欧姆接触区附近复合,因此比起向下注入的电子,横向注入电子的输运系数相当低。另外由于 ϕ_i 随掺杂浓度的升高而升高(等式(4.2.10)),那么在给定偏压下 ϕ_i 大的地方电子注入减少,因此横向注入通常被自然地抑制。另外由于横向基区没有任何附加电场,这也能抑制横向注入。设计者可以通过浅发射区、大的横向基区尺寸或者用一个很窄的氧化隔离区将发射极和基极接触分开等方法减小横向电子注入。

集成电路晶体管的表面积比一般的集成电路电阻和电容小,因此在设计电路的时候要尽可能使用晶体管,只有在必要时才加入其他器件。如果用双极晶体管作为有源器件,那么它们通常是集成电路所有器件中的主要部分。尽管晶体管的表面积略大于第 5 章提到的简单的 pn 结二极管,但通常还是用晶体管代替二极管,这是因为用不同连接方式的晶体管等效的二极管具有较低的串联电阻和较小的开关时间。共有五种方法可以将晶体管连接成二极管,因为只需要连接三端器件中的两端。不同的二极管连接方式如图 6.23 所示,其中最常用的是图 6.23(a)中的连接方式,因为它串联电阻最小,开关速度最快[17]。

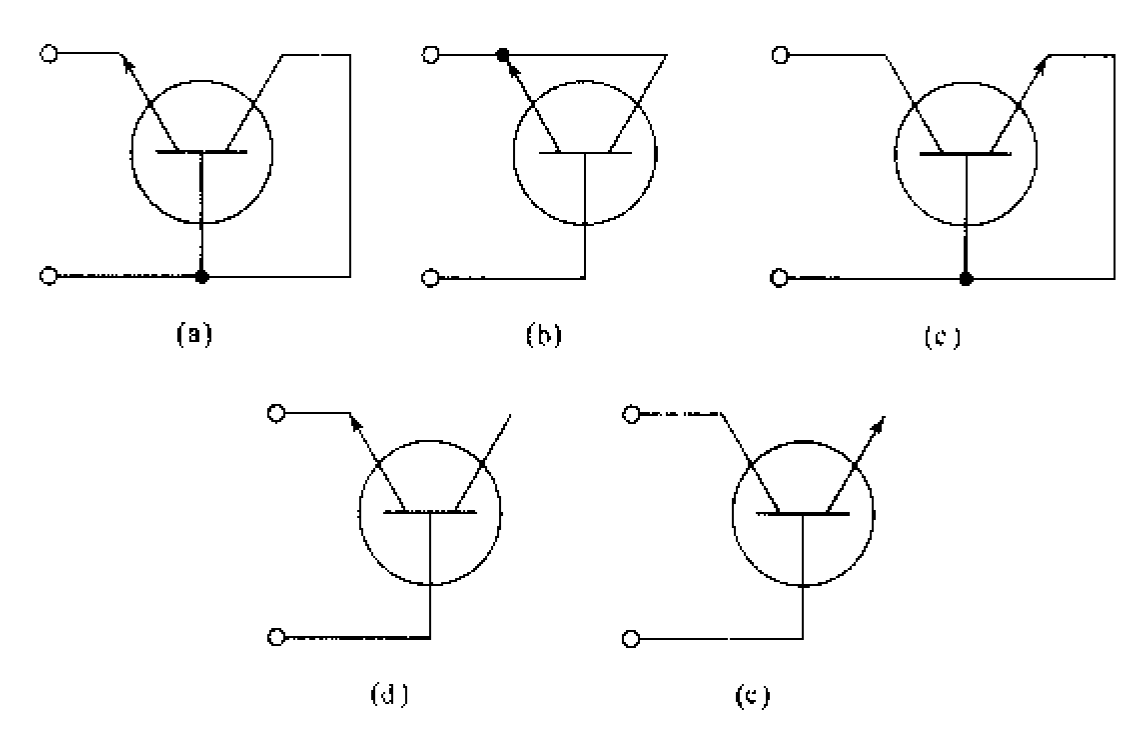


图 6.23 将晶体管改接成二极管的五种方式

6.6 器件:异质结双极晶体管

由 6.2 节知,双极晶体管的增益受到发射效率 γ 的限制。对于 npn 晶体管,γ 取决于发射区注入到基区的电子电流 J_n(有用成分)与基区注入到发射区的空穴电流 J_p(无用成分)的比值(等式(6.2.20))。降低基区掺杂浓度和减薄基区可以提高发射区和基区 Gummel 数的比值,从而提高增益和频率响应。然而,高性能的电路同时需要高的增益、高的频率响应(7.6 节)和低的基区电阻。如果能够避免等式(6.2.20)中的基区电阻和增益之间的折衷,晶体管和电路的性能将得到很好的改善。由第 5 章知,穿过同质 pn 结的电子电流与空穴电流的比值约等于发射区与基区 Gummel 数的比值,但异质结中电子和空穴的注入比还包含一项附加因子 $\exp(\Delta E_g/kT)$。我们可以在异质结双极晶体管 HBT 中利用这个附加因子[10,11]。

4.2 节和 5.3 节讨论了结两侧分别为不同的半导体材料的单个 pn 异质结的性质,我们将以此为基础讨论 HBT。假设晶体管的发射结为异质结,其附近的另一个结可以是同质结也可

以是异质结。如4.2节和5.3节介绍的,考虑n型发射区的带隙大于p型基区的带隙:$E_{gE}>E_{gB}$。首先假设基区和集电区由同种材料组成,如图6.24(a)所示,这种晶体管称为单异质结双极晶体管(基区和集电区也可以由不同材料组成,即图6.24(b)所示的双异质结双极晶体管,我们将在本节的后半部分讨论)。下面先分析在单异质结晶体管和双异质结晶体管中结构相似的发射结。

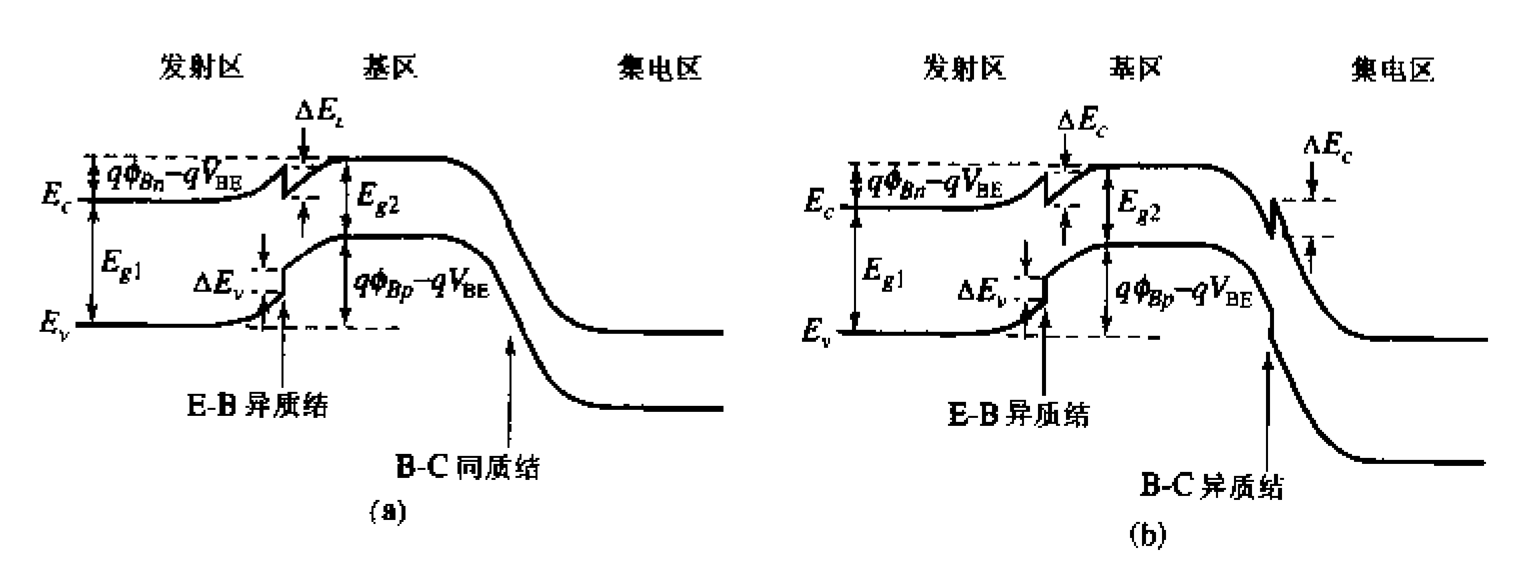


图6.24 异质结双极晶体管的能带图:(a)单异质结双极晶体管;(b)双异质结双极晶体管

如图6.24(a)所示,异质结上基区导带边的弯曲程度大于导带的不连续。由于准中性基区的导带边高于发射结异质结处的导带尖峰,因此扩散电流起主导作用。由等式(5.3.35)和(5.3.36)可知,当扩散电流为主时,比值J_n/J_p随$\exp(\Delta E_g/kT)$的升高而升高。这样,如果晶体管的增益受到发射效率的限制,那么增益随两种材料总的带隙差指数上升。一般来说,异质结双极晶体管的增益要比同质结双极晶体管大几个数量级。由于发射效率中含有包含带隙差的指数因子,使得即使在基区掺杂浓度(和Gummel数GN_B)较高、基区电阻较低的情况下,仍能保证足够的增益。

随着基区掺杂浓度的上升或发射区掺杂浓度的下降,发射区能带弯曲增大而基区能带弯曲减小。当基区能带弯曲减小至小于ΔE_C时,异质结上的尖峰部分高于基区导带边,如图5.11(c)所示。在这种情况下,电子注入电流J_n不再是扩散电流,而是穿越异质结导带尖峰的热电子发射电流,并且势垒ϕ_{Bn}大于等式(5.3.30)中的值。空穴注入电流J_p仍然可以由等式(5.3.31)中给定的势垒ϕ_{Bp}描述,因此比值J_n/J_p和增益都将降低。当基区重掺杂时,所有的能带弯曲都发生在发射区内,$\phi_{Bn}=X_B-X_E=\Delta E_c$,$\phi_{Bn}=\phi_i$;$\phi_{Bp}=(\phi_i+X_E+E_{gE})-(X_B+E_{gB})=\phi_i-\Delta E_c+\Delta E_g$。电子和空穴电流的比值取决于$\phi_{Bp}-\phi_{Bn}$,在这种情况下,$\phi_{Bp}-\phi_{Bn}=\Delta E_g-\Delta E_c=\Delta E_v$,不再直接与带隙差$\Delta E_g$有关,因而无法得到HBT的全部优点。因为增益与能量成指数关系,所以导带尖峰升高时增益会急剧下降。另外,如图5.11(d)所示,小电流时扩散电流占主导成分。但随着外加偏压的升高,基区的导带被拉至低于导带尖峰,此时,热电子发射电流占主导地位。

这样,异质结双极晶体管的发射效率为

$$\gamma=\left[1+\frac{GN_B}{GN_E}\exp\left(\frac{-\Delta E_x}{kT}\right)\right]^{-1} \tag{6.6.1}$$

这里$\Delta E_x=\Delta E_g$或者ΔE_v,与发射结异质结的尖峰和基区导带边的相对高低有关。

例题 突变异质结双极晶体管

最常见的异质结双极晶体管的基区是 GaAs,宽带隙发射区是 AlAs 与 GaAs 的化合物($Al_xGa_{1-x}As$),通常铝组分约为30%,镓组分约为70%($x=0.3$)。AlAs 与 GaAs 的带隙以及其他性质随化合物中组分比例的变化而缓慢变化,但不完全线性[12]。

(a) 首先考虑基区和发射区中等掺杂时的能带图:$N_{dE}=2\times10^{17}\ \text{cm}^{-3}$,$N_{aB}=5\times10^{16}\ \text{cm}^{-3}$。此时基区的导带边高于异质结处的导带尖峰,由等式(6.6.1)与材料的性质,可以计算出室温下的晶体管增益(为简化数学计算,假设电子和空穴的扩散系数相同,发射区和基区的厚度相等,则有 $GN_B/GN_E=N_{aB}/N_{dE}$)。

$$E_g(\text{GaAs}) = 1.424\ \text{eV}$$
$$E_g(\text{AlAs}) = 2.168\ \text{eV}$$
$$E_g(\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}) = 1.424 + 1.247 \times 0.3 = 1.798\ \text{eV}\ [12]$$
$$\Delta E_g = 1.247 \times 0.3 = 0.374\ \text{eV}$$
$$\exp(\Delta E_g/kT) = \exp(0.374/0.0259) = 1.87 \times 10^6$$
$$GN_B/GN_E = 0.25$$
$$\gamma = (1 + 0.25/1.87 \times 10^6)^{-1} = (1 + 1.34 \times 10^{-7})^{-1}$$
$$\beta = 7.5 \times 10^6$$

然而,当增益只受注入载流子的比值限制时才能得到如此大的β值。实际上,除了发射效率外还有其他因素影响增益。发射结空间电荷区内的复合很可能发生在化合物半导体系统的界面,并且中性基区内的复合也能影响增益。

如果上述的晶体管不是异质结结构,则发射效率为0.8(也就是说反向空穴注入将占总电流的20%),增益约等于4。

(b) 将基区掺杂浓度提高到$1\times10^{19}\ \text{cm}^{-3}$来减小基区电阻,异质结处的导带尖峰将高于基区导带边,如图5.11(c)所示。此时所有的电压都降落在发射区,γ不再与ΔE_g有关,而是依赖于ΔE_v。

$$X(\text{GaAs}) = 4.07\ \text{V}$$
$$X(\text{AlAs}) = 3.5\ \text{V}$$
$$X(\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}) = 4.07 - 1.1 \times 0.3 = 3.74\ \text{V}\ [12]$$
$$\Delta E_c = 0.33\ \text{eV}$$
$$\Delta E_v = 0.044\ \text{eV}$$
$$\exp(\Delta E_v/kT) = \exp(0.044/0.0259) = 5.5$$
$$GN_B/GN_E = 50$$
$$\gamma = (1 + 0.25 \times 50)^{-1} = (1 + 12.5)^{-1}$$
$$\beta < 1$$

这种掺杂浓度下的晶体管是不实用的。

为减小或消除导带的尖峰效应,可以在发射结空间电荷区沿厚度方向使组分从纯 GaAs 向 $Al_xGa_{1-x}As$ 渐变。这种方法以更复杂的材料结构为代价提高了器件的性能。如果缓变的异质结材料使导带尖峰减小至零,那么载流子注入要越过的势垒将由体材料的性质决定,并且发射效率与ΔE_g有关,可以再次用含有总带隙差ΔE_g的等式(6.6.1)来求发射效率

$$\Delta E_g = 1.247 \times 0.3 = 0.374\ \text{eV}$$
$$\exp(\Delta E_g/kT) = \exp(0.374/0.0259) = 1.87 \times 10^6$$
$$GN_B/GN_E = 50$$
$$\gamma = (1 + 50/1.87 \times 10^6)^{-1} = (1 + 2.67 \times 10^{-5})^{-1}$$
$$\beta \approx 3.7 \times 10^4$$

同样,其他因素的影响可能会限制增益,使其低于这个理想值。但是即使基区高掺杂,仍能获得相当可观的增益。

6.6.1 双异质结双极晶体管

前面分析突变异质结双极晶体管时,主要讨论的是关键的发射结。如果晶体管的集电结也是异质结,那么还要考虑一些其他效应。此外,实际情况下的基区带隙也小于集电区。如果集电结的价带有显著的不连续,集电区的空穴注入和存储都会下降,这对工作在饱和区的晶体管影响很大。工作在放大区时,如果导带不连续,将会阻碍基区向集电区输运电子,如图6.25(a)所示。在某些实际电路中,集电结上的高反偏电压能够减小这个势垒。缓变集电结也有利于降低势垒,提高集电结的电子收集效果。如果基区杂质从窄带隙的基区扩散到宽带隙的集电区,将在界面处形成一个电势势垒,如图6.25(b)所示,也会阻止电子从基区到集电区的输运。

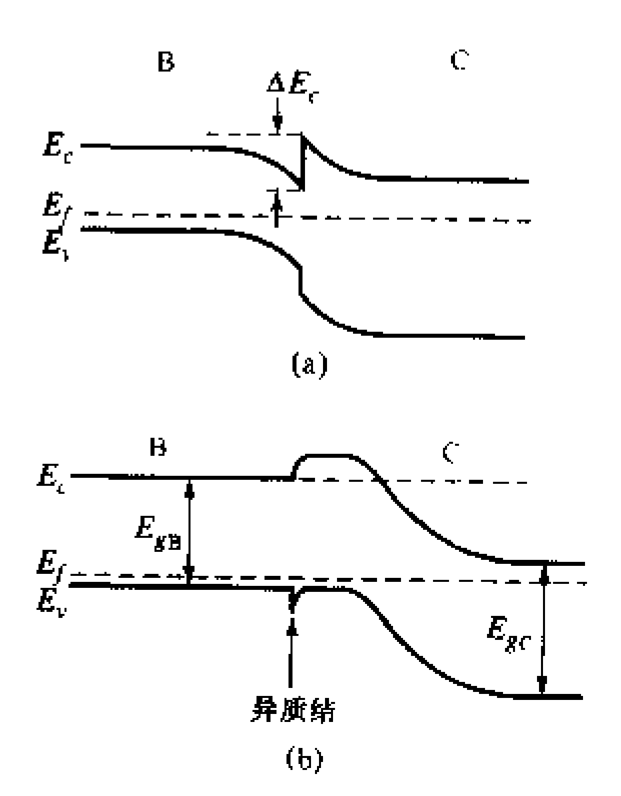


图6.25 (a)集电结异质结上的导带不连续形成了势垒,阻止电子从基区进入到集电区;(b)(杂质扩散或大电流效应引起的)中性基区的扩展也能形成阻止电子被集电区收集的势垒

在7.2节我们将看到,大电流情况下,基区内的过剩自由电子会扩大中性基区的范围,进而增加异质结双极晶体管的渡越时间和空穴存储,这与同质结晶体管类似。不过,在这种情况下异质结晶体管的器件性能还会有其他退化。如果中性基区扩展从窄带隙材料进入宽带隙材料,在集电结处将形成一个阻止电子流动的势垒,如图6.25(b)所示。准中性基区内的 $E_f - E_v$ 为常数,由此我们可以了解这个势垒的形成原因。当带隙变大时,为了保证 $E_f - E_v$ 为常数、导

带边必须上翘,形成了阻止电子流向集电极的势垒,降低了晶体管的增益和频率响应。在某些情况下,基极-集电极偏压可以降低这个附加势垒,但异质集电结还是需要仔细设计。

6.6.2 准中性基区的带隙缓变

通过改变准中性基区的半导体材料组分可进一步优化器件性能。准中性基区的带隙缓变是非常实用的,能够在基区形成电场、提高电子速度、降低电子在基区的渡越时间。

为了了解电场的起源及其优点,我们继续研究 4.1 节和图 4.2(b) 中提到的带隙缓变的概念,并且重新给出图 6.26。文献[10]中指出,作用在半导体中的电子和空穴上的力等于载流子所处的带边的斜率(不考虑电子符号),在同质结构中带隙为常数,因此导带和价带边的斜率相等,作用在电子和空穴上的力大小相等方向相反;在异质结结构中,带隙随位置变化,因此导带边和价带边的斜率可以不同,作用在电子和空穴上的力也可以不同。实际上,除了电场外,异质结还利用带隙变化来控制电子和空穴的分布与流动。改变带隙为器件设计提供了一个新的途径,尤其是在电子和空穴的分布与流动都能被控制的器件。

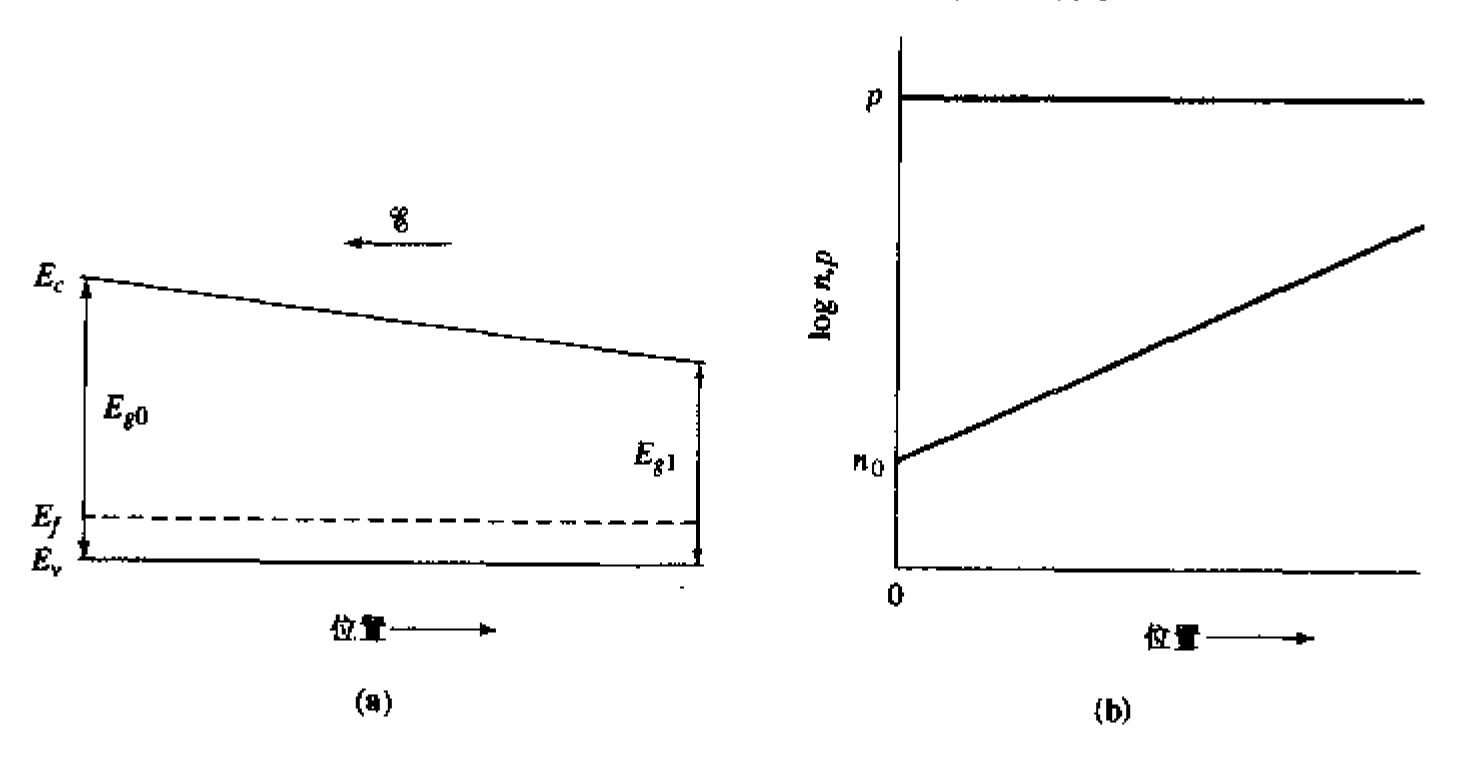


图 6.26　(a)带隙随位置变化的 p 型材料的能带图;(b)相应的载流子浓度分布

就基区掺杂浓度高的晶体管来说,空穴浓度与 $E_f - E_v$ 一样大约为常数,缓变效应出现在导带边,如图 6.26 所示。作用在电子上的电场为 $dE_c/dx \approx dE_g/dx$。

文献[13]对双极晶体管基区中的带隙缓变做了详细分析,这里总结一下结论。关键问题是材料特性,为描述方便,可以用与位置相关的本征载流子浓度来表示。利用本章前面推导出的常规表达式,并考虑与位置的关系,重新推导得到小注入情况下(小电流或高基区掺杂)的电子电流为[13]

$$J_n = -\frac{q\exp\left(\frac{qV_{BE}}{kT}\right)}{\int_0^{W_B}\frac{N_{ab}}{D_n n_i^2}dx} \tag{6.6.2}$$

等式(6.6.2)与电子电流的一般表达式(等式(6.1.17),在下面重新给出)相似,只是现在 n_i 是位置的函数。

$$J_n = \frac{-qD_n n_{i0}^2}{N_{ab}W_B}\exp\left(\frac{qV_{BE}}{kT}\right) \tag{6.6.3}$$

因为带隙升高时 n_i 急剧下降，基区中带隙最大的部分对等式(6.6.2)分母中积分的贡献最大。

还有一个很重要的特殊情况，基区带隙从发射结边缘到集电结边缘线性下降

$$n_i^2 = n_{i0}^2 \exp\left(\frac{\Delta E_g}{kT}\right)\frac{x}{W_B} \tag{6.6.4}$$

将等式(6.6.4)代入等式(6.6.2)得到电子电流。考虑到带隙的变化远大于 kT，那么

$$J_n = \frac{-qD_n n_{i0}^2}{N_{ab}W_B}\frac{\Delta E_g}{kT}\exp\left(\frac{qV_{BE}}{kT}\right) \tag{6.6.5}$$

比较等式(6.6.5)与基区带隙均匀时的电子电流表达式(等式(6.6.3))，发现带隙渐变使 J_n 随因子 $\Delta E_g/kT$ 的升高而增大。还可以将等式(6.6.5)写成与内建电场 $\mathscr{E}$ 相关的形式

$$J_n = \frac{-q\mu_n n_{i0}^2}{N_{ab}}\mathscr{E}\exp\left(\frac{qV_{BE}}{kT}\right) \tag{6.6.6}$$

因为空穴电流 J_p 随带隙缓变影响不大，所以增益随 $\Delta E_g/kT$ 的升高而增加。虽然增益的增加非常有用，但电子通过基区输运时间的降低更为重要。我们将在7.6节研究基区渡越时间问题。

例题　均匀带隙 HBT 和缓变带隙 HBT

本例比较 $Si/Si_{1-x}Ge_x$ 异质结均匀带隙和缓变带隙晶体管。均匀带隙 HBT 的掺杂分布如图6.27(a)所示。在两层硅之间插入一层锗含量为常数的 $Si_{1-x}Ge_x$ 区。正如2.6节提到的，埋层和中等掺杂的集电区用外延淀积得到，接着用气相外延淀积 $Si_{1-x}Ge_x$ 层，反应气体同时含有锗源和硅源，然后在 $Si_{1-x}Ge_x$ 上淀积一层硅形成发射区。这三层外延层通常依次连续淀积，这样衬底不会暴露在空气中形成氧化层。

大部分锗硅层中重掺杂了p型硼来形成基区，而没有被掺杂的锗硅缓冲层包围重掺杂的基区，防止在淀积了 $Si_{1-x}Ge_x$ 层后的热处理过程中硼扩散到锗硅层之外。在集电极一侧，这种隔离尤为重要。前面提到，大电流情况下可动载流子能够将中性基区的边缘延伸至集电区，形成阻止电子电流流向集电区的势垒，降低了晶体管性能。因此，p型基区一定要被控制在窄带隙材料中。

采用异质发射结的 HBT 具有前面提到的各种优点，尤其是能够减少基区向发射区反向注入的空穴，这样就可以在保持高发射效率和高增益的同时，提高基区掺杂浓度以减小基区电阻。发射区的掺杂浓度甚至可以低于基区掺杂浓度，如图6.27(a)所示。发射区表面通常为重掺杂以减小发射极接触电阻。

图6.27(b)中的基区带隙缓变 HBT 与基区带隙均匀的 HBT 相比，基区中的锗含量从发射结边缘到集电结边缘逐渐升高。集电结为突变异质结，其设计考虑与基区带隙均匀的 HBT 相同。然而发射结附近的情况与集电结极的不同，从基区向发射区注入的反向空穴势垒只是稍高于同质结，因此为保持发射效率不能提高基区掺杂浓度。基区带隙缓变最大的优点就是形成了一个电场，能够降低基区渡越时间。这个问题将在7.6节中继续讨论。图6.28为典型的掺杂浓度和锗含量剖面图。

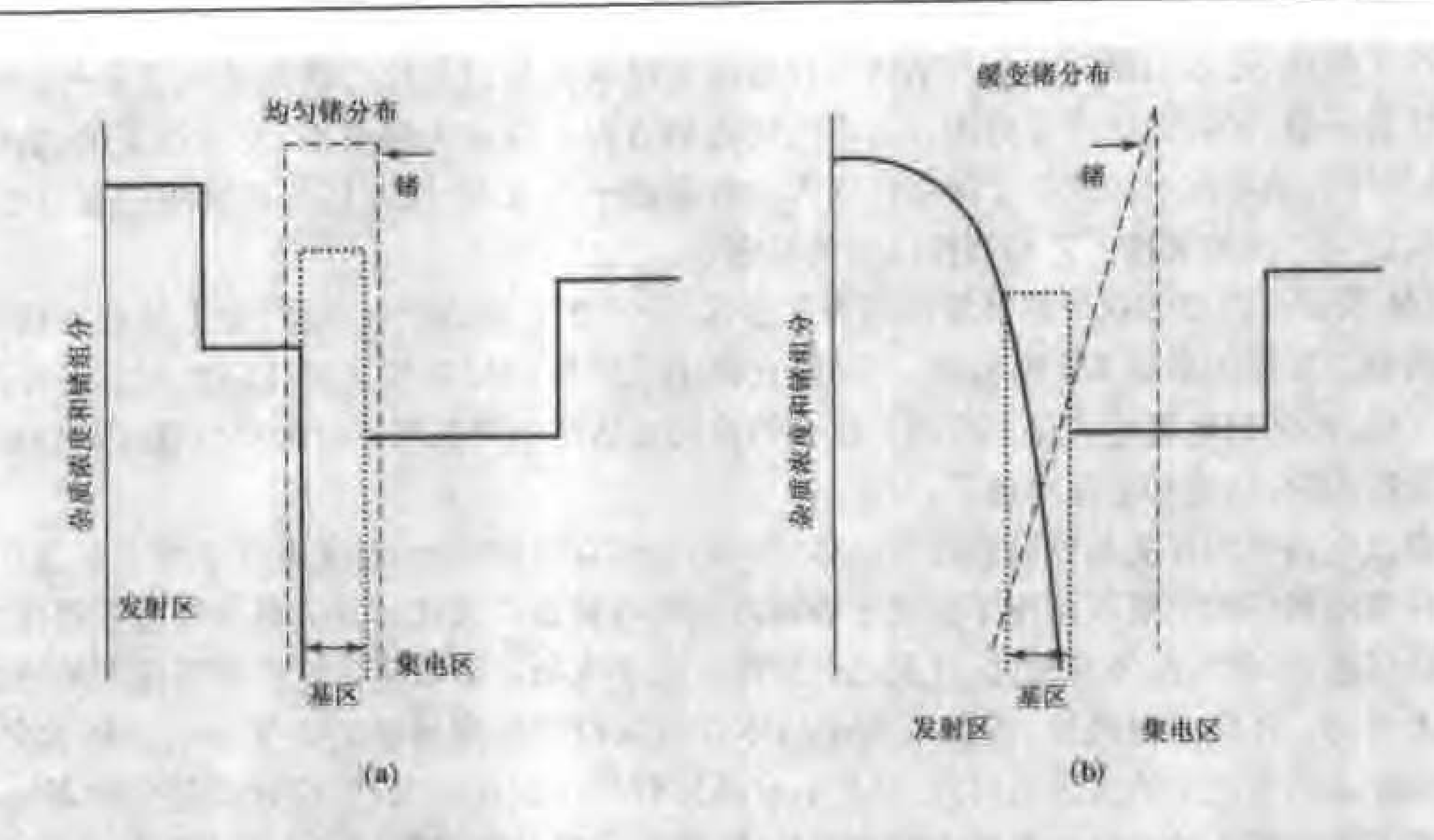


图 6.27　以硅(发射区)-$Si_{1-x}Ge_x$(基区)-硅(集电区)HBT 为例的均匀带隙(a)和缓变带隙;(b)异质结双极晶体管

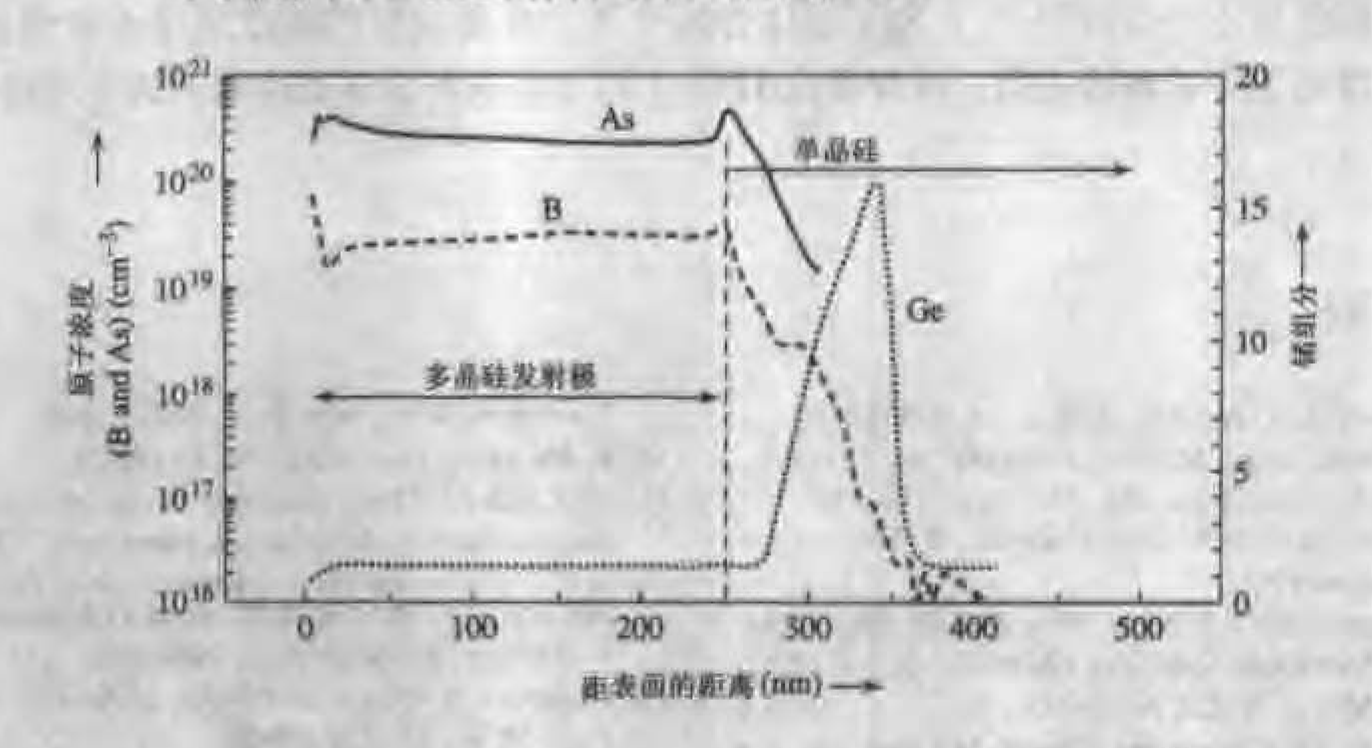


图 6.28　缓变带隙 $Si/Si_{1-x}Ge_x$ HBT 中锗、砷和硼含量的剖面图,其中多晶硅接触区位于单晶硅发射区的上方[18] (© 1994 IEEE)

小结

我们可以通过研究非平衡 pn 结少数载流子的行为来理解双极晶体管的工作原理。pn 结的空间电荷区能够阻止多数载流子流动,收集少数载流子,因而当两个 pn 结靠得很近且背靠背放置时,一个结上的偏压就能够影响另一个结附近少数载流子的数目,显著改变其电学性能。这就是双极晶体管效应,利用该效应可以制备有效的开关和放大器件。

晶体管基区中沿垂直两个 pn 结的方向上多数载流子电流近似为零,因为多子在该方向上总会遇到一个势垒,由此可以推导出描述晶体管特性的方程。两个 pn 结之间的电流主要由少

数载流子构成,连接电流的大小与基区多数载流子电荷总量成反比。晶体管电流是外加电压的非线性函数,实际设计中可利用各工作区的近似方程。在放大偏置下,发射结正偏,集电结反偏,晶体管的电流增益依赖于向基区注入少数载流子的效率,同时还与基区中的复合有关。这些效应与器件的制造工艺和器件结构关系密切。

EM 模型可以描述双极晶体管的所有工作区,是非常重要的一级模型,也是其他更精确模型的基础。互易关系是 EM 模型的一部分,它指出晶体管的反向放大和正向放大只是外部电压的互换,而输出电流等式是不变的。这个结论同时适用于对称和不对称晶体管,因为电流增益下降的同时,结饱和电流增加了。

集成电路中的开关晶体管和放大晶体管一般制作在外延层上,且集电区下方有埋层存在。这两种实用器件的主要区别在于垂直于表面方向的各种器件尺寸。开关晶体管通常制作在薄的外延层上,与放大晶体管相比,开关晶体管需要减小电荷存储效应。这两种器件所采用的工艺也不相同。在集成电路中,传统的 npn 晶体管被以衬底为集电极的寄生 pnp 晶体管分流。虽然 npn 晶体管处于放大区的时候,寄生 pnp 晶体管处于截止状态,但在开关过程中寄生 pnp 晶体管有可能进入放大区。在双极集成电路中,因为双极晶体管的尺寸小于其他器件,所以电路设计者会尽可能地利用双极晶体管来代替二极管或电阻。异质结双极晶体管利用了不同半导体之间的能带不连续,进一步控制了晶体管的性能。异质发射结可以增加电流增益或在相同的电流增益下降低基区电阻。缓变基区 HBT 中的基区电场能够加速电子通过基区,减小基区渡越时间。

参考文献

1. W. Shockley, *Bell Syst. Tech. J.*, **28**, 435 (1949).
2. J. L. Moll and I. M. Ross, *Proc. IRE*, **44**, 72 (1956).
3. H. K. Gummel, *Proc. IRE*, **49**, 834 (1961).
4. H. Kroemer, *Arch. Elek. Übertrag.*, **8**, May, August, November (1954).
5. J. J. Ebers and J. L. Moll, *Proc. IRE*, **42**, 1761 (1954).
6. J. W. Slotboom, *Solid-State Electronics* **20**, 279 (1977).
7. C. S. Meyer, D. K. Lynn and D. J. Hamilton, *Analysis and Design of Integrated Circuits*, McGraw-Hill, New York, 1968, pp. 248–258.
8. H. Camenzind, *Electronic Integrated Systems Design*, Copyright 1972 by Litton Educational Publishing, Inc. Reprinted by permission of Van Nostrand Reinhold Company.
9. P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. F. Gibbons, *Physical Electronics and Circuit Models of Transistors*, Wiley, New York, 1964, p. 145.
10. H. Kroemer, *Proc. IEEE*. **70**, 13 (1982).
11. H. Kroemer, "Heterostructure device physics: Band discontinuities as device design parameters," Chapter 4 in *VLSI Electronics Microstructure Science*. **10**: *Surface and Interface Effects in VLSI*, eds. N. G. Einspruch and R. S Bauer, Academic Press, New York, 1985.
12. Properties of AlGaAs and GaAs: S. Adachi, *J. Appl. Phys.* **58**, pp. R1–R29 (1985).
13. H. Kroemer, *Solid-State Electronics* **28**, 1101 (1985).
14. T. Kamins, *Polycrystalline Silicon for Integrated Circuits and Displays*, Second Edition, Kluwer Academic Publishers, Boston, 1998, p. 272, Figure 6.10.
15. D. Vook, et al., IEEE Trans. Electron Devices, **41**, 1013 (1994)

参考书

S. M. Sze, *Semiconductor Devices: Physics and Technology*, second edition, Wiley, New York, 2002.

B. G. Streetman and S. Banerjee, *Solid-State Electronic Devices*, fifth edition, Prentice-Hall, Upper Saddle River, NJ (2000).

A Sedra and K. Smith, *Microelectronic Circuits*, fourth edition, Oxford Univ. Press, 1998.

A. G. Milnes, *Semiconductor Devices and Integrated Electronics*, Van Nostrand Reinhold, New York, 1980.

习题

6.1* 等式(6.1.2)指出,掺杂浓度随位置指数变化的区域中存在恒定电场。如果晶体管基区存在 −4000 V/cm的恒定电场,基区宽度为0.3μm,发射结空间电荷区边缘的基区掺杂浓度为 $10^{17}cm^{-3}$,求集电结空间电荷区边缘的基区掺杂浓度。

6.2 将等式(6.1.10)应用到放大偏置的 npn 晶体管,设 $I_C = -J_n A_E$,证明

$$I_C = \frac{qA_E\tilde{D}_n N_a(x)n(x)}{\int_x^{x_B} N_a d\xi}$$

6.3* 考虑两种 npn 晶体管:(i)习题6.1 中的晶体管;(ii)除了基区掺杂浓度恒定为 $10^{17}cm^{-3}$外,其余条件与习题6.1 相同。两种晶体管均为放大偏置,并且集电极电流相同,利用习题6.2 中的表达式求出两种晶体管的基区电子浓度与 x 的关系,并在同一个坐标系下绘出 $n(x)$随 x 的变化曲线。解释为什么两图中 $n(x)$ 在集电结空间电荷区边缘处具有相同的梯度。

6.4* 推导习题6.3 中描述的两种晶体管基区内储存的少子电荷总量的表达式。假设基极电流只包括基区内的复合电流,利用上面得到的结果比较两个器件的 β_F值。

6.5* (a)利用图6.4 中的晶体管数据,求准中性基区中的多子电荷 Q_{BO}。

(b)如果基区平均掺杂浓度为 $10^{17}cm^{-3}$(假设 $\tilde{D}_n = 25cm^2s^{-1}$),求基区宽度 x_B。

6.6* 近似求解图6.15 中的开关晶体管和图6.16 中的放大晶体管的 Gummel 数,并与习题6.5 中晶体管的 Gummel 数进行比较,说明结果的差异。

6.7 假设发射结和集电结都是线性缓变结,已知浓度梯度 $a = 10^{22}cm^{-4}$(其中 $N_d - N_a = ax$),$\phi_i = 0.872V$,重新求解6.3 节中的例题。

6.8* 下表给出了净掺杂浓度($N_d - N_a$)随位置的变化,晶体管的掺杂浓度分布如图6.16 所示。

净掺杂浓度($N_d - N_a$)(cm^{-3})	距发射结的距离(μm)
7.36×10^{17}	−0.125
-1.25×10^{16}	+0.045
-4.0×10^{16}	+0.115
-1.9×10^{16}	+0.15
-6.4×10^{15}	+0.205
-1.5×10^{15}	+0.250

假设晶体管的有源区面积为 $2\times10^{-6}cm^2$。(a)在线性坐标系中画出掺杂分布,说明这个分布图可以用 $N_a \approx 2\times10^{16}cm^{-3}$的单边突变结大致近似。(b)作出与图6.11 类似的 Q_{VE}随总的结电压变化的曲线,结电压的偏置范围为 +0.3V 到击穿电压。求出如图6.11 所示 ϕ_i和 K_t的值。

6.9 晶体管的掺杂分布如图6.16 所示,用习题6.1 的结果和习题6.8 的数据估算准中性基区中电场的大小。

6.10 画出与图6.12 类似的 pnp 晶体管的等效电路图,写出与等式(6.4.9)与(6.4.10)类似的表达式来描述 pnp 晶体管的 EM 模型。

6.11 证明可用图6.14(b)描述由基极电流驱动的晶体管 Ebers-Moss 模型。

6.12* 根据 EM 模型推导晶体管饱和压降表达式6.4.17。已知 $I_C/I_B = 10$,$\alpha_F = 0.985$,$\alpha_R = 0.72$,求 $V_{CE(sat)}$。

6.13 I_{CE0}是基极开路时流入反偏集电结的电流,I_{CB0}是发射极开路时流入反偏集电结的电流。用pnp晶体管的EM方程(习题6.10)求出这两个电流的比值,并从物理上解释二者的差异。

6.14[†] 考虑有源偏置下的npn晶体管,$t=0$时刻,一束强光照射到集电结空间电荷区,单位时间在空间电荷区间产生G个空穴-电子对(qG大约与直流基极电流I_B属同一量级)。

(a)如果发射极-基极电压和集电极-基极电压都为常数,那么$t>0$时的基极、发射极和集电极电流为多少?

(b)如果基极由电流源驱动,光照不能改变I_B,重新求解(a)中的问题。

6.15 在室温附近,处于放大偏置的npn晶体管的集电极电流I_C和基极电流I_B通常为正值。假设一个npn晶体管处于放大偏置,并且随温度升高I_C恒为常数。如果测量I_B发现其数值下降并且最终变成负值,那么可以用什么物理效应来解释这一现象?(考虑所有的基极电流分量)[9]

6.16 已知npn晶体管的基区和发射区都是非均匀掺杂,且反向注入发射区的空穴在发射极电极接触处复合,证明发射效率γ可以写成$\gamma=(1+Q_{B0}\bar{D}_{pE}/Q_{E0}\bar{D}_{nB})^{-1}$,其中$Q_{E0}$为准中性发射区的多子电荷总量,$Q_{B0}$为准中性基区多子电荷总量,$\bar{D}_{pE}$和$\bar{D}_{nB}$为这些区域的少数载流子有效扩散系数。证明在相同的复合条件下这个结果与原型晶体管给出的结果一致(等式(6.6.20))。

6.17[†] 已知npn晶体管的横截面积为10^{-5}cm^2,基区均匀掺杂且$N_a=4\times10^{17}\text{cm}^{-3}$,$D_{nB}=18\text{cm}^2\text{s}^{-1}$,$x_B=0.5\mu\text{m}$。

(a)如果发射区的条件与习题6.16类似,发射区的杂质总量(Q_{E0}/q)为8×10^9个原子,$\bar{D}_{pE}$为$2\text{cm}^2\text{s}^{-1}$,计算$\gamma$。

(b)如果基区寿命为10^{-6}s,计算基区输运系数α_T。

(c)计算这个晶体管的β_F。如果用近似表达式$\beta_F\approx Q_{E0}\bar{D}_{nB}/Q_{B0}\bar{D}_{pE}$,误差是多少?这个近似表达式有时会用来估算$\beta_F$。

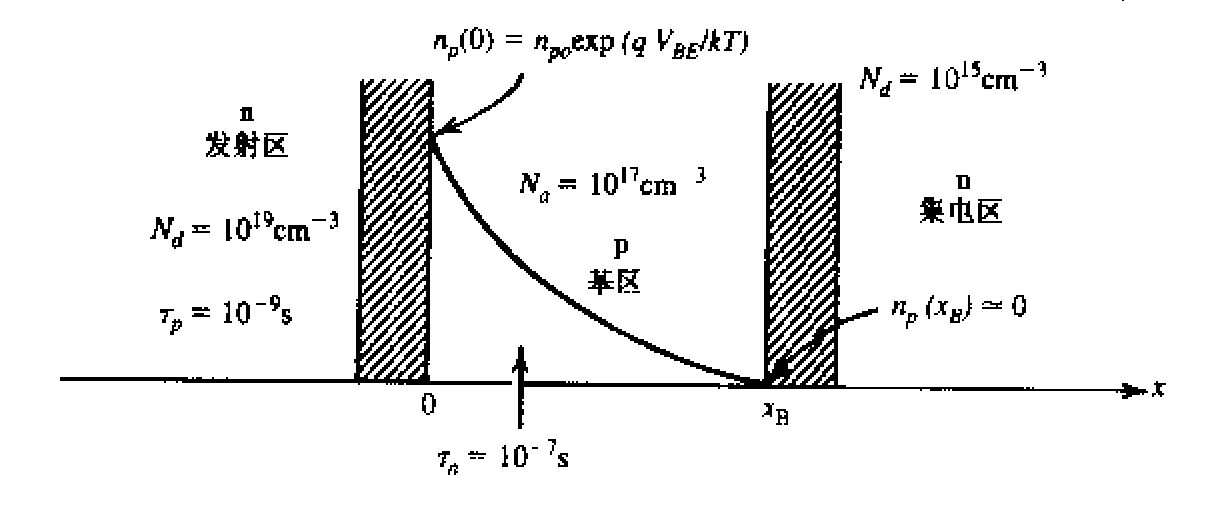
6.18[†] 假设习题6.17中的晶体管被放置在辐射环境中,电子寿命以$\tau_n=\tau_{n0}\exp-(t/t_d)$规律衰减。该等式中的$t$每天测量一次,$t_d=10$天,画出$\beta_F$随天数变化的曲线,并求出$\beta_F$下降至1时所需要的时间,以及此时的基区寿命。

6.19[†] 晶体管结构如图P6.19所示。

(a)假设基区存在复合,推导基区中电子浓度与x的关系。(需要解一个类似等式(5.3.10)的方程)

(b)$x=0$处的电子浓度梯度与注入电子电流成正比,$x=x_B$处的电子浓度梯度与被收集的电子电流成正比,二者的差异来源于基区复合导致的电流损失(由基极电极提供),基区复合电流是L_n和x_B/L_n的函数[$L_n=(D_n\tau_n)^{1/2}$],写出基区复合电流的表达式。

(c)假设均匀掺杂的发射区、基区和集电区的掺杂浓度分别为10^{19}cm^{-3}、10^{17}cm^{-3}和10^{15}cm^{-3},发射区空穴寿命为10^{-9}s,基区电子寿命为10^{-7}s,发射区和基区的宽度都为$1\mu\text{m}$,计算发射效率和β_F。利用图1.16查出扩散系数。[提示:这道题最好将连续性方程表示成双曲线函数$\cosh(x/L_n)$与$\sinh(x/L_n)$的和]。



图P6.19

6.20[†]　如果用电子浓度 n 和电子运动速度 v_d 来表示集电极电流($I_C = qnv_dA$),利用已知的放大工作的npn原型晶体管基区内的注入电子分布,求基区渡越时间 τ_B。也就是将渡越时间表示为 $\tau_B = \int_0^{x_B} dx/v$,可得到 $\tau_B = x_B^2/2D_n$。

第7章 双极晶体管Ⅱ:局限性与模型化

迄今为止,我们已经讨论了双极晶体管基本的工作原理。例如,偏置在放大区的晶体管的工作原理可以这样描述:发射结具有较高的注入效率,从发射区注入到基区的载流子大大增加了基区的少子浓度。这些非平衡少子在基区中从发射结向集电结扩散,当到达集电结空间电荷区边界时,被电场扫入准中性集电区,形成集电区多子电流,流向集电极。

在实际的晶体管中,这种简单的描述在较大的范围内是比较准确的。然而,我们还需要考虑非理想情况。例如,第6章描述的理想晶体管在放大区的输出为电流源输出,也就是说,输出电流与基极和集电极之间的外加电压无关。然而,实际晶体管的输出电流和输出电压都会影响到器件特性。又如,理想晶体管电流增益的理论预示值与偏置条件无关,但这只是一个近似。本章讨论的第一个问题是实际晶体管相对于理想晶体管的附加效应。

本章要解决的第二个问题是双极晶体管的动态特性。使用电荷控制模型将大大简化动态特性的分析,并可得出双极晶体管的大信号模型。将EM模型和电荷控制模型结合起来将得到一种非常有效的,被广泛用于晶体管计算机模拟的分析方法。从大信号模型还可直接推导出小信号等效电路,即混合Π电路。由于混合Π等效电路中的元件与晶体管的物理关系比较直接,因此目前被普遍采用。电荷控制的概念还有助于从物理上理解双极晶体管的高频工作极限。

本章的最后将介绍在标准的IC工艺中制造pnp双极晶体管的两种典型方法:衬底pnp晶体管和横向pnp晶体管。衬底pnp晶体管由于集电区不是独立的,其应用受到了限制。横向pnp晶体管的性能也远不如npn晶体管。我们将通过简单的模型来解释这些问题。

7.1 基区宽度调变效应(Early效应)

第6章分析放大偏置的晶体管时,只考虑了集电极-基极电压的基本作用:保证集电区对基区少子的有效收集,并将其输运至集电极。第6章中没有考虑集电极-基极反偏电压对集电极电流的影响,反偏电压大小只是受到击穿电压的限制。然而,我们在第4章已经指出,反偏pn结的空间电荷区宽度明显依赖于偏置电压。这种依赖关系是场效应晶体管的工作基础,也使得作为线性放大器的晶体管的特性变得复杂。集电极-基极偏置电压的变化将引起集电结空间电荷区宽度的改变,导致准中性基区宽度变化,从而影响到晶体管的特性,这种现象称为基区宽度调变效应。James Early[1]最早分析了这种现象,所以也称之为Early效应。

集电极电流与集电极-基极电压的关系可以直接利用6.2节中放大偏置npn晶体管的公式得出。利用等式(6.2.1)和(6.1.15),得到

$$I_C = \frac{q\tilde{D}_n n_i^2 A_E \exp(qV_{BE}/kT)}{\int_0^{x_B} p \, dx} \tag{7.1.1}$$

等式中各项的定义见 6.1 节，其中 x_B 是 V_{CB} 的函数。

V_{CB} 的改变对集电极电流的影响可表示为

$$\frac{\partial I_C}{\partial V_{CB}} = \frac{-q\tilde{D}_n n_i^2 A_E \exp(qV_{BE}/kT) p(x_B)}{\left[\int_0^{x_B} p \, dx\right]^2} \frac{\partial x_B}{\partial V_{CB}} \tag{7.1.2}$$

等式(7.1.2)表示的是集电结的小信号电导，等式右端的几项组合在一起可以表示集电极电流本身，因此可以将该等式写成

$$\begin{aligned} \frac{\partial I_C}{\partial V_{CB}} &= -I_C p(x_B)\left[\frac{1}{\int_0^{x_B} p \, dx}\right]\left[\frac{\partial x_B}{\partial V_{CB}}\right] \\ &= -\frac{I_C}{V_A} = \frac{I_C}{|V_A|} \end{aligned} \tag{7.1.3}$$

由于集电结是反向偏置的，等式(7.1.3)中的 $\partial x_B/\partial V_{CB} < 0$，所以 Early 效应将导致 I_C 随 V_{CB} 的上升而增加。从晶体管的共发射极输出特性曲线，如图 7.1(a)所示，可明显看出这一效应。图 7.1(a)的横坐标是 V_{CE}，但对于偏置在放大区的晶体管，V_{CE} 的变化类似于 V_{CB} 的变化（$V_{CE} \approx V_{CB} + 0.7\text{V}$）。

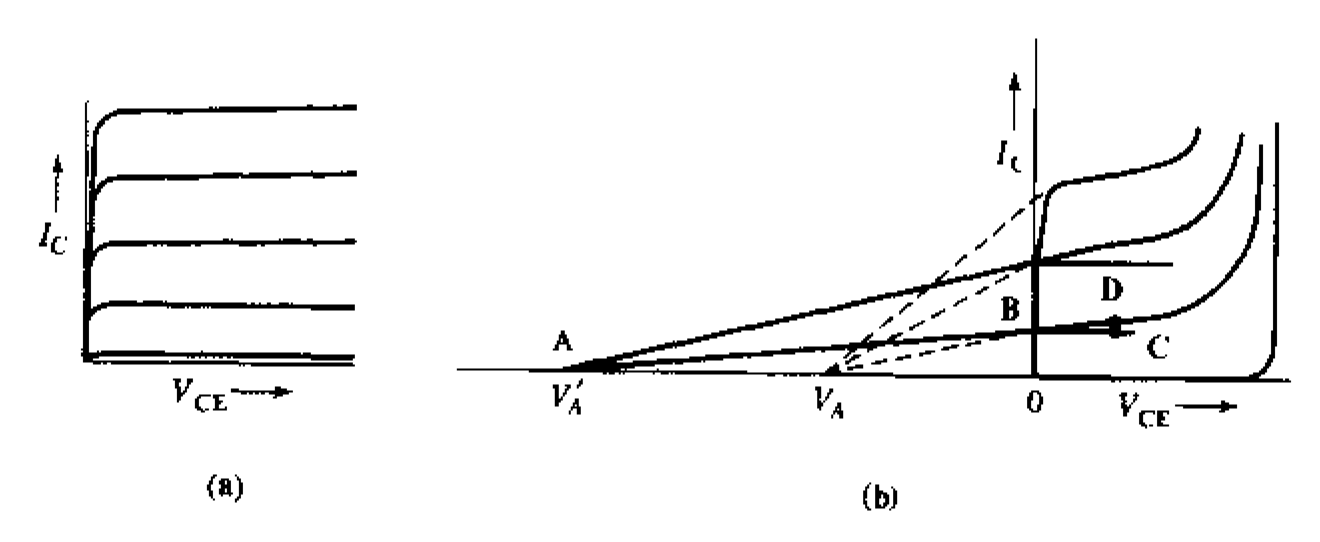


图 7.1　工作在放大区的晶体管的共发射极输出特性测量曲线。(a)纵轴 0～0.5mA，横轴 0～5V；(b)纵轴 0～0.5mA，横轴范围为 50V，测试曲线起点（饱和区边缘）处的切线（虚线）延长到电压轴可确定 Early 电压 V_A。放大区切线的延长线（实线）与电压轴的交点为 V'_A

等式(7.1.3)表明 Early 效应与集电极电流之间为线性关系，比例因子的倒数乘以电流 I_C 具有电压的量纲，该电压定义为 Early 电压，用 V_A 表示。从等式(7.1.3)可得到 npn 晶体管的 Early 电压的定义为

$$V_A = \frac{\int_0^{x_B} p \, dx}{p(x_B) \partial x_B / \partial V_{CB}} \tag{7.1.4}$$

由等式(7.1.4)得出的V_A小于零,所以npn晶体管的Early电压为负值。工作在反向放大区的晶体管,由于发射结空间电荷区宽度的改变而导致的基区宽度调变效应,对应的Early电压用V_B表示。

如果不考虑大电流效应,V_A定义中的三项因素只取决于晶体管的制造工艺和集电极-基极电压。实际上,集电极-基极电压对V_A本身的影响通常不予考虑,因此Early电压可用某一偏置(通常取$V_{CB}=0$)下的值近似。在这种情况下,绘出$I_C \sim V_{CB}$(实际上用的是V_{CE})曲线的正向放大区边界处(V_{CB}近似为零)的切线,其延长线与横轴的交点就是Early电压V_A(图7.1(b)中的虚线)。对于电路设计和分析来说,我们感兴趣的不是临界饱和点的Early电压,而是正向放大区的Early电压。如果在放大区画出$I_C \sim V_{CE}$曲线的切线,它们往往在电压轴上不会交于同一点。然而,这些切线在晶体管的偏置范围内可近似相交于电压轴上的V_A',如图7.1(b)中的实线所示。

为了得到V_A的更为实用的表达式,我们对等式(7.1.4)进行整理。首先,利用等式(6.1.13)将分子表示成单位面积基区多子电荷Q_B的函数

$$\int_0^{x_B} p\,dx = \frac{Q_B}{q} \tag{7.1.5}$$

等式(7.1.4)中的分母可表示成Q_B对V_{CB}微分

$$qp(x_B)\frac{dx_B}{dV_{CB}} = \frac{dQ_B}{dV_{CB}} \tag{7.1.6}$$

等式(7.1.6)右边的微分项可与单位面积集电结小信号电容C_μ对应

$$\left|\frac{dQ_B}{dV_{CB}}\right| = C_{jc} \tag{7.1.7}$$

这样,Early电压表示成

$$|V_A| = \frac{Q_B}{C_{jc}} \tag{7.1.8}$$

为减小集电极-基极电压对集电极电流的影响,应增大Early电压V_A的数值,由等式(7.1.8)知,应增大单位面积基区多子电荷总量与单位面积集电结电容的比值。从物理上看,这将减小集电结空间电荷区向基区的扩展。

晶体管模型中用Early电压V_A描述基区宽度调变效应,该内容将在7.7节介绍。我们可以利用第6章的原型晶体管从物理概念上来理解Early效应。如图7.2所示,工作在放大区的原型晶体管,基区少子分布为三角形。集电极-基极电压从V_{CB}增大到$(V_{CB}+\Delta V_{CB})$时,集电结空间电荷区边界从x_{B0}移动到x_{B1},基区少子分布也随之改变。图中阴影部分反映了基区存储电荷的减少量。

因此,V_{CB}的增大将导致集电极电流按x_{B0}/x_{B1}比例增加。分析基区存储电荷的变化,我们还会发现Early效应减小了基区电荷的存储量,这会影响晶体管的瞬态特性和基极电流,因为基区复合直接与基区存储的电荷量有关。这些效应将在7.5节晶体管模型中详细讨论。

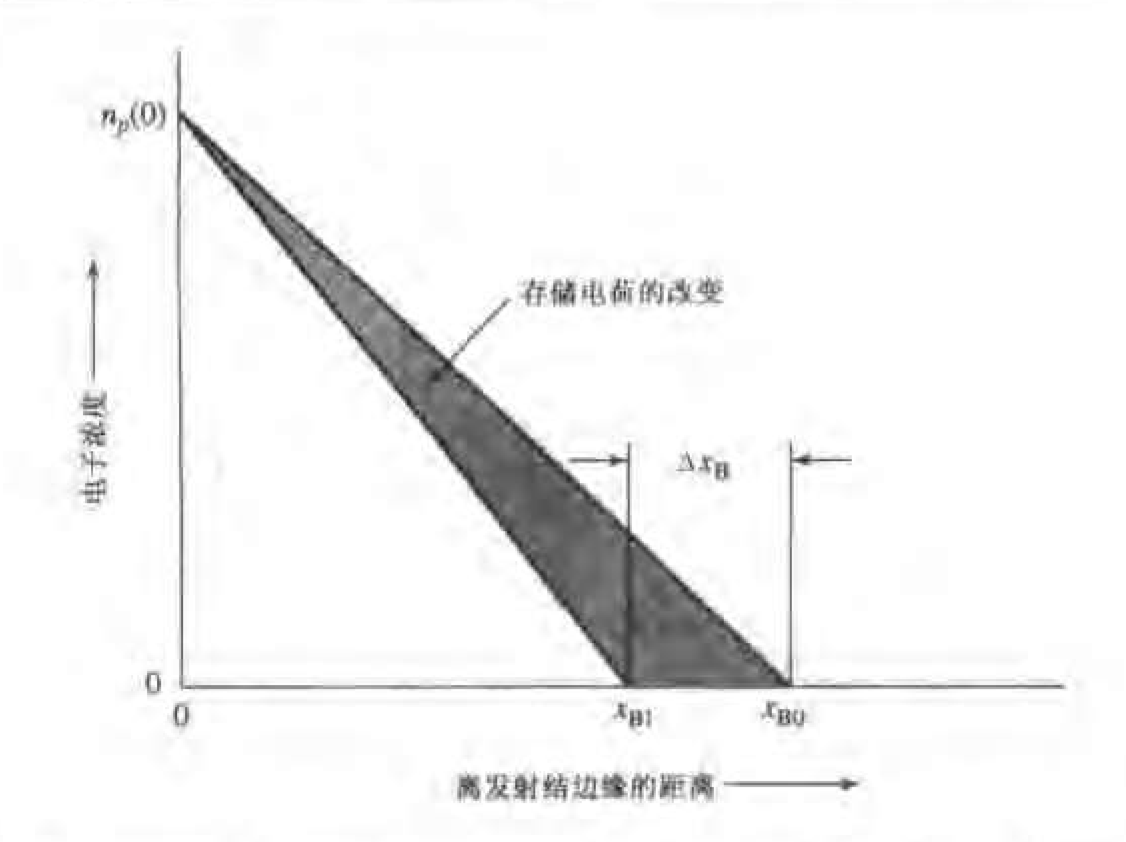


图7.2 Early效应对准中性基区少子分布的影响。V_{CB}的增加减小了基区宽度，使基区边界从x_{B0}移动到x_{B1}，从而增大了n_p的梯度，减小了基区少子电荷总量

7.2 发射结低偏置效应和高偏置效应

通过研究放大区工作的晶体管的基极电流I_B、集电极电流I_C与发射极电压V_{BE}之间的关系，我们会发现一些偏离理想特性的现象。考虑到晶体管的特点，在半对数坐标系中分析电流电压关系更为直观。图7.3是典型的IC中用于放大电路的npn晶体管测量曲线。从图中可以看出，在中等电流区，I_C和I_B的对数与V_{BE}的关系近似为直线，这与第6章的分析相吻合。但在发射结高偏置和低偏置时，电流电压关系曲线将偏离理想特性，其中集电极电流I_C符合理想特性的范围要大于基极电流I_B的理想范围。本节将讨论这些非理想现象，我们首先讨论在很低的偏置电压下，I_B偏离理想特性的问题。

7.2.1 发射结低偏置电流

发射结低偏置时，$\log I_B \sim V_{BE}$曲线的斜率下降。实验数据表明，随着V_{BE}下降至零，基极电流与电压V_{BE}的关系趋于等式(7.2.1)

$$I_B = I_0 \exp\left(\frac{qV_{BE}}{nkT}\right) \tag{7.2.1}$$

式中n通常取1和2之间的值，I_0的数值比中等偏置区相应的指数项的系数大。

从第5章pn结电流的分析中我们很容易知道，在低偏置下基极电流偏离理想特性是因为发射结空间电荷区的复合引起的额外的发射结电流。5.3节利用Shochley-Hall-Read理论推导出了二极管空间电荷区复合电流的表达式(5.3.24)，该表达式与电压的关系与等式(7.2.1)类似，如果n=2则相同。n的取值在1和2之间改变反映了空间电荷区复合电流对基极电流影响的程度。等式(5.3.25)指出，随着偏置电压的减小，复合电流与注入到准中性区的电流相比，其

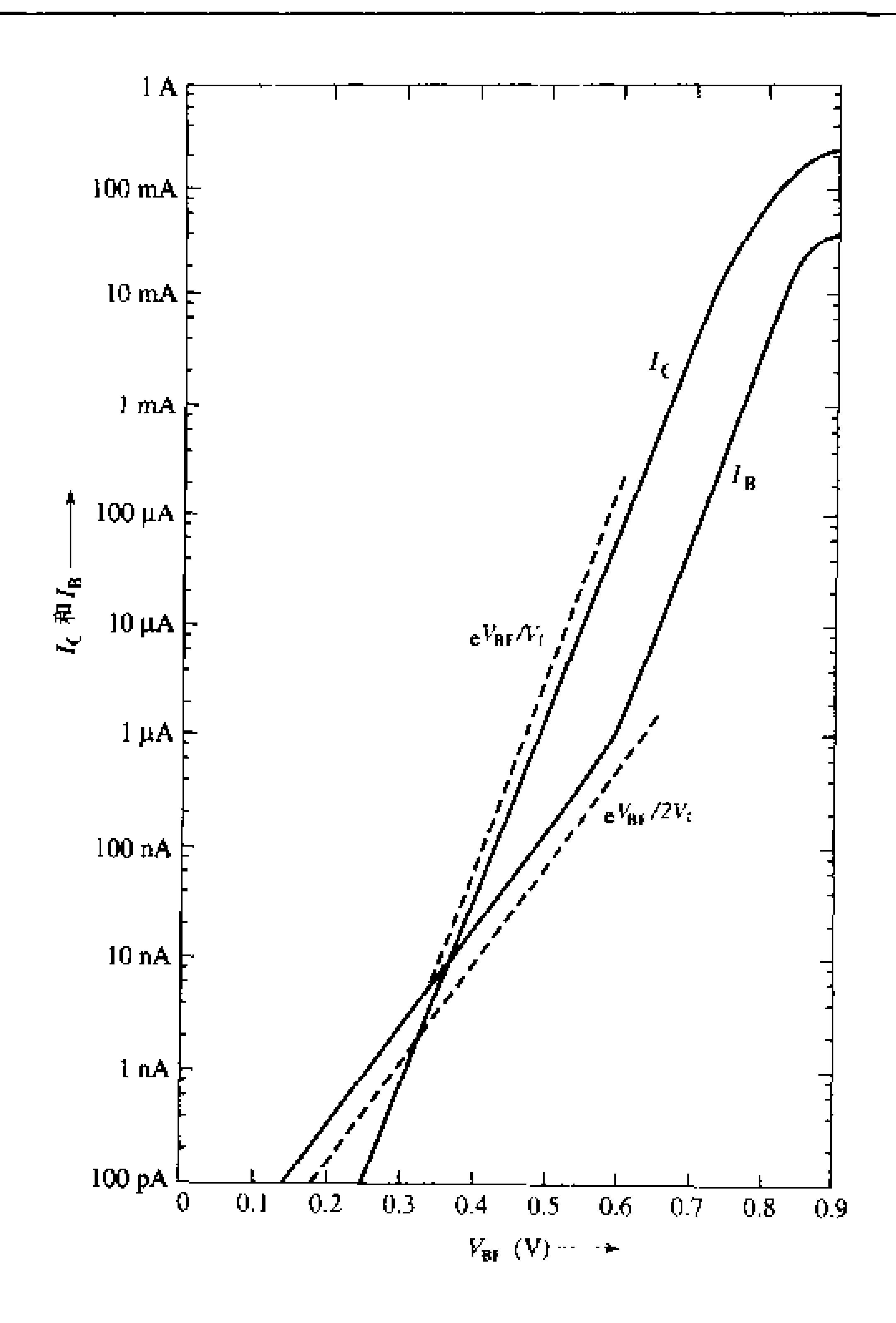


图 7.3 典型的工作于正向放大区的晶体管，集电极电流 I_C 和基极电流 I_B 与基极-发射极电压 V_{BE} 的关系

重要性增加。例如，由对等式(5.3.25)的讨论知，当结电压低于 0.48V 时，复合电流起主要作用

发射结空间电荷区复合电流只会影响基极和发射极电流，不会影响到集电极电流，因此，集电极电流与基极-发射极电压的关系仍由等式(7.1.1)决定。随着 V_{BE} 的减小，当集电极电流的主要成分由基区扩散流转变为集电结空间电荷区产生流时，集电极电流改由等式(6.3.1)描述。因此，发射结低偏置区的集电极电流在发射极电流中所占比例要比中等偏置区低，导致电流增益下降。该结论可以通过 $\beta_F \sim I_C$ 曲线(图 7.4)得到验证。图 7.4 是利用等式(6.2.22)从图 7.3 得到的。低偏置区 β_F 的下降限制了晶体管用于放大电路的低电压工作范围。

维持晶体管低偏置区的高电流增益，减小空间电荷区的复合电流具有非常重要的商业意义，例如医用的助听器、起搏器电路都需要在尽可能低的电流下维持足够好的性能。用于此类功能的电路，在器件制造过程中应尽可能提高发射结空间电荷区的载流子寿命。

7.2.2 大注入

迄今为止，对晶体管的理论分析是在小注入近似下进行的，即外加偏置时准中性区的多子

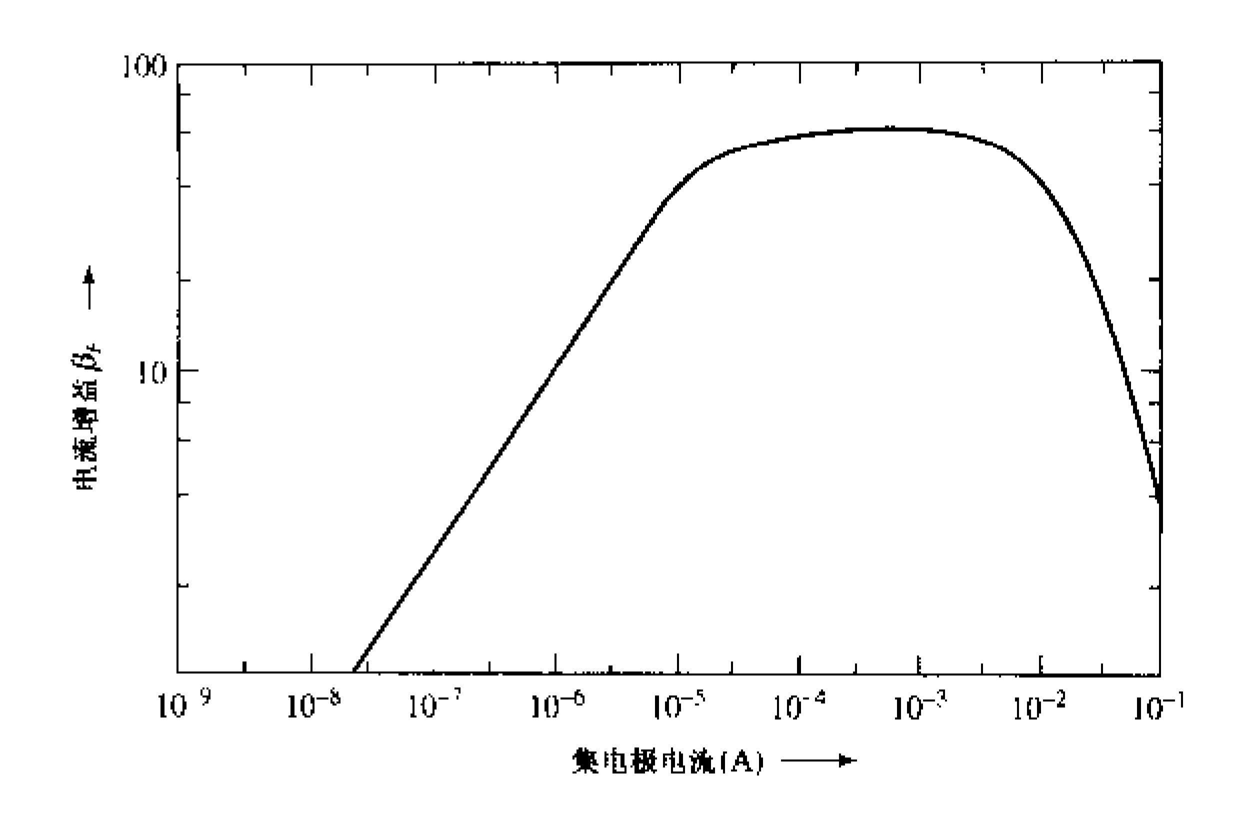


图7.4 对于图7.3的晶体管,电流增益β_F与集电极电流I_C的关系

浓度近似等于其热平衡值。当工作在放大区的晶体管的发射结偏置提高时,在注入少子浓度高或者本身掺杂浓度低的地方,都会出现偏离小注入近似的情形。第一种情形出现在基区中靠近发射结的位置,第二种情形出现在IC晶体管的集电结附近。实用的晶体管这两种情形都有可能出现。下面我们分别讨论。

发射结的大注入效应 偏置在放大区的晶体管的集电极电流基本表达式(7.1.1)是由等式(6.1.12)推导出来的。图7.3验证了等式(7.1.1)预示的指数关系,并指出这种指数关系在集电极电流变化的大约八个数量级的范围内是有效的。发射结高偏置时该电流电压关系将偏离该指数关系。

导致偏离的原因之一是发射区向基区的大注入效应。如果注入到基区的电子浓度足够高,导致基区的多子浓度显著提高,那么电流电压关系将明显偏离等式(7.1.1)。在这种情形下,基区的多子浓度跟偏置有关(相关的二级效应是扩散系数$\bar{D}_n$也会改变)。考虑到基区的准中性,基区多子的积分为

$$\int_0^{x_B} p(x)\mathrm{d}x = \int_0^{x_B}[N_a(x) + n'(x)]\mathrm{d}x \tag{7.2.2}$$

式中$n'(x)$为注入的电子浓度,$N_a(x)$为基区掺杂浓度。进一步推导需要确定积分中的项与位置的关系,因此要考虑基区的杂质分布。

然而,定性分析只需要重点考虑基区边界处的注入电子浓度$n'(0)$,因为该处电子浓度最高,直接影响到积分值的大小。利用等式(5.3.20)和准中性近似可以求出$n'(0)$。等式(5.3.20)指出了发射结边界处的pn乘积与偏置电压之间的关系。当电流密度远低于热平衡时的电流趋势(热平衡时的漂移电流或扩散电流,5.3节估计该值为$\sim 10^5\text{A cm}^{-2}$)时,关系式$p(0)n(0) = n_i^2\exp(qV_{BE}/kT)$依然有效[1],因为该量级的电流不会破坏pn结的热平衡。

放大偏置时基区边界少子浓度$n(0)$近似等于边界处注入的电子浓度$n'(0)$

---

1 使用晶体管的基本表达式(6.1.12)时,也需要等式(5.3.20)在大注入时成立,因为等式(6.1.12)也是利用等式(5.3.20)推导出来的。

$$n(0)=\frac{N_a(0)}{2}\left[\left(1+\frac{4n_i^2\exp(qV_{BE}/kT)}{N_a^2(0)}\right)^{1/2}-1\right] \tag{7.2.3}$$

小注入以及中等注入时,等式(7.2.3)可简化成等式(5.3.7)。大注入时,平方根中的第二项起主要作用,$n(0)$将趋于与$\exp(qV_{BE}/2kT)$成正比。大注入时的集电极电流也趋于与$\exp(qV_{BE}/2kT)$成正比。

集电结的大注入效应(Kirk效应) 在介绍小注入的概念时已经指出,掺杂浓度越低,越不容易满足小注入条件。典型的IC晶体管(见图6.16)中,掺杂浓度最低的区域是靠近基区的集电区,此处的外延材料中没有附加掺杂。采用低掺杂集电区的目的是:提高集电结击穿电压,降低集电结电容,减小集电结空间电荷区向基区的扩展(抑制了Early效应)。本节将分析低掺杂集电区带来的一些不利影响。

偏置在放大区的晶体管,由于集电结是反偏的,任何到达集电结边界的少子都会被很快地扫入另一侧,因此可以假定集电结空间电荷区边界少子浓度为零。但实际上,这个假定只是一个近似,只要有载流子通过空间电荷区,就会违反该假定。这是因为考虑到自由载流子的速度极限,大小为J的电流密度要求少数载流子浓度至少等于J/qv_l,其中v_l为饱和漂移速度(等式(1.2.12))。这些电子在中性基区可被空穴(多子)中和,但在集电结空间电荷区将与电离杂质原子提供的固定电荷叠加,改变了空间电荷分布。当注入的电荷浓度增加到与掺杂浓度相当时,必须用与集电极电流有关的电荷分量修正前面的理论。随着集电极电流的增加,这一机制将显著改变晶体管的特性。

该效应可以这样定性描述:通过集电结空间电荷区的自由载流子改变该区原有的电荷分布,从而改变电场分布。如果固定V_{CB},则空间电荷区电场的积分为常数,那么空间电荷浓度的改变将导致空间电荷区宽度改变。我们将看到,随着集电极电流的增大,基区一侧的集电结空间电荷区减薄,中性基区宽度增大,这种现象称为Kirk效应[3]。

图7.5(a)为图6.16所示晶体管的局部杂质分布图。如果集电极-基极电压$V_{CB}=10\text{V}$,不考虑自由载流子电荷对集电结空间电荷区的影响,那么该晶体管的集电结电场分布如图7.5(b)所示。图中的$\mathscr{E}_l$为极限散射临界电场强度。当电场强度超过$\mathscr{E}_l$时,自由电子将以饱和漂移速度v_l漂移(见图1.18)。考虑了自由载流子电荷影响的npn晶体管集电结空间电荷区的Poisson方程为

$$\frac{d\mathscr{E}}{dx}=\frac{1}{\epsilon_s}\left[qN(x)-\frac{J_C}{v(x)}\right] \tag{7.2.4}$$

式中$N(x)=N_d-N_a$是空间电荷区的净掺杂浓度(基区一侧为负的电离受主,高电阻外延层和重掺杂埋层为正的电离施主)。固定V_{CB}不变,可得到包含电场$\mathscr{E}$的表达式

$$V_{CB}+\phi_i=\int_{x_p}^{x_n}-\mathscr{E}\,dx \tag{7.2.5}$$

式中ϕ_i为pn结内建电势(等式(4.2.10))。等式(7.2.5)隐含了对空间电荷区宽度的描述。再次考虑等式(7.2.4),我们注意到等式右边第二项在小电流时可以忽略,大电流时必须考虑。令括号中的两项相等,可得到划分小电流区和大电流区的临界电流密度J_1

$$J_1=qN(x)v(x) \tag{7.2.6}$$

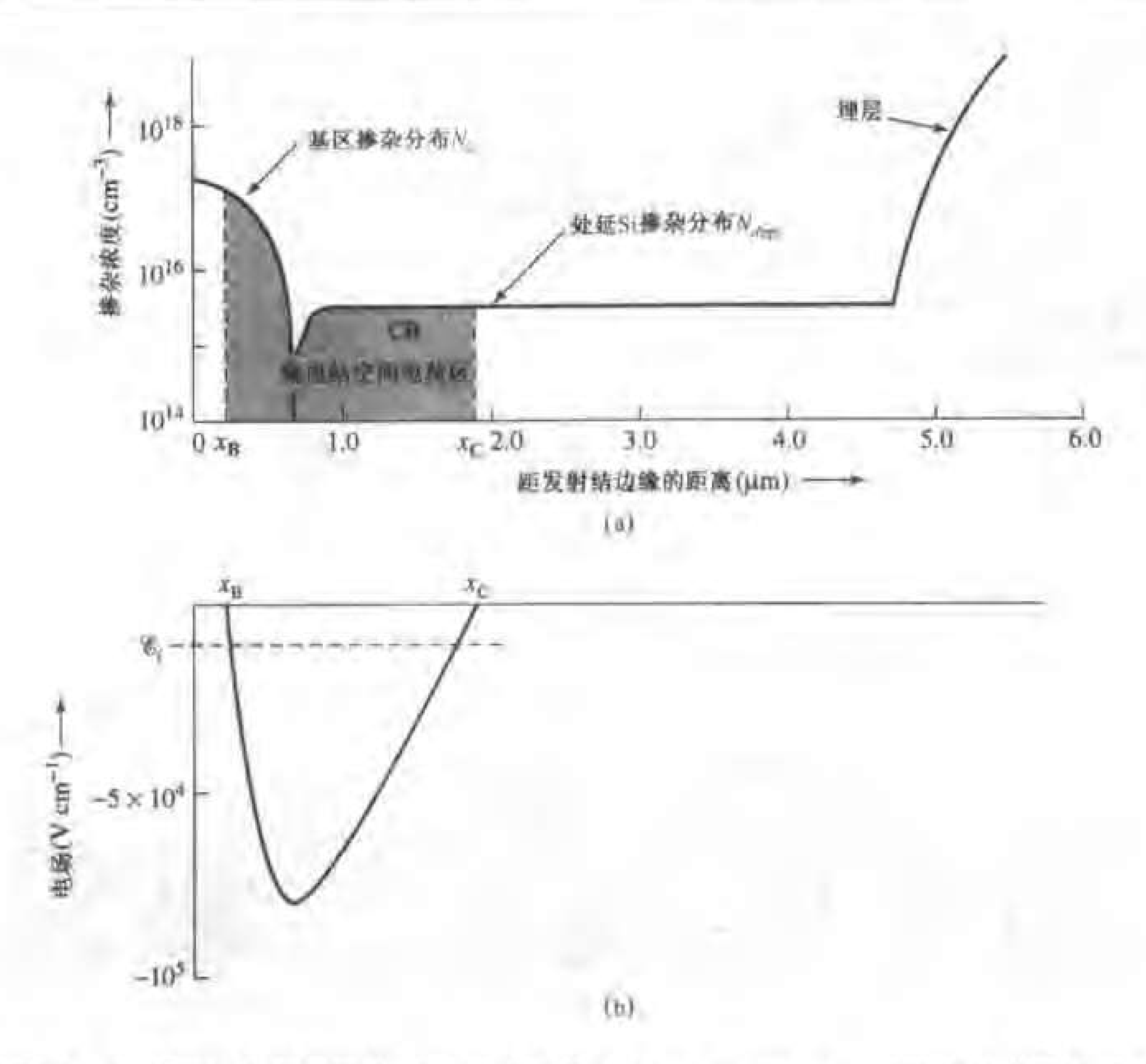


图7.5　对于图6.16的晶体管，工作于正向放大区，$V_{CB}=10V$。(a)掺杂分布和集电结空间电荷区分布；(b)集电结空间电荷区电场分布，图中$\mathscr{E}_l$为极限散射临界电场强度

由于外延层净掺杂浓度$N(x)$最小，所以外延层的临界电流密度最先达到。考虑到外延层中集电结空间电荷区大部分范围内的载流子漂移速度为v_l，外延层掺杂浓度$N(x)=N_{epi}$，所以$J_1=qN_{epi}v_l$。该效应的分析有重要的实际意义，因为在许多应用中，晶体管的集电极电流密度等于或大于临界电流密度J_1。例如，如果$A_E\approx10^{-7}\text{cm}^2$，$N_{epi}=5\times10^{15}\text{cm}^{-3}$，那么临界集电极电流为0.8mA。

接下来进行数学推导，首先在等式(7.2.4)两端各乘x，然后在集电结空间电荷区，即x_B和x_C之间积分

$$\int_{x_B}^{x_C}x\frac{\mathrm{d}\mathscr{E}}{\mathrm{d}x}\mathrm{d}x=\frac{1}{\epsilon_s}\int_{x_B}^{x_C}x\left[qN(x)-\frac{J_C}{v(x)}\right]\mathrm{d}x \tag{7.2.7}$$

等式(7.2.7)的左边可分步积分并利用等式(7.2.5)求出

$$\int_{x_B}^{x_C}x\,\mathrm{d}\mathscr{E}=-\int_{x_B}^{x_C}\mathscr{E}\,\mathrm{d}x=V_{CB}+\phi_i \tag{7.2.8}$$

其中取$\mathscr{E}(x_C)=\mathscr{E}(x_B)=0$。那么，集电极-基极电压可写成

$$V_{CB}=\frac{1}{\epsilon_s}\int_{x_B}^{x_C}x\left[qN(x)-\frac{J_C}{v(x)}\right]\mathrm{d}x-\phi_i \tag{7.2.9}$$

由于杂质的非均匀分布，用等式(7.2.9)求解图7.5不是很直观。为了避免复杂的数学分析，并能从物理上更为清楚地解释Kirk效应，我们利用图7.6所示的并不常用的外延基区晶体管进行分析。该器件用高电阻率p型外延层替代了常见的n型外延层。对于做在p型外延层上的npn晶体管，其集电区是高掺杂的n^+埋层，集电结空间电荷区边界x_C的位置几乎不受V_{CB}的影响，因此可以假设x_C为常数。利用等式(7.2.9)就可以得到空间电荷区宽度$x_{CB}=x_C-x_B$与集电极电流之间的关系。由于受主杂质均匀分布，我们取外延层的$qN(x)$为常数$-qN_{epi}$，直至埋层位置，并假设$v=v_l$，因此

$$\begin{aligned} V_{CB} &= \frac{1}{2\epsilon_s}\left[qN_{epi}+\frac{J_C}{v_l}\right]x_{CB}^2-\phi_i \\ &= \frac{qN_{epi}}{2\epsilon_s}\left[1+\frac{J_C}{J_1}\right]x_{CB}^2-\phi_i \end{aligned} \tag{7.2.10}$$

利用上式可求出集电结空间电荷区宽度x_{CB}。

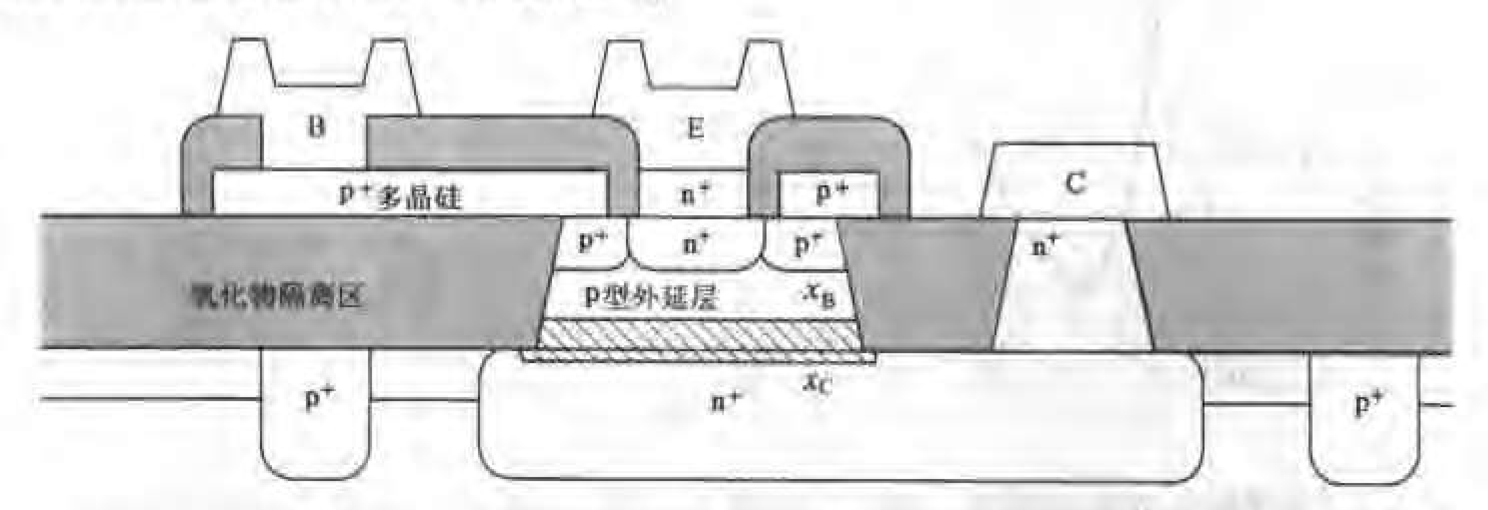


图7.6 外延基区npn晶体管横截面示意图(氧化物隔离，多晶硅基极和发射极)。图中斜线阴影部分代表集电结空间电荷区的扩展范围(x_B到x_C)

集电极电流等于零时的集电结空间电荷区宽度x_{CO}为

$$x_{CO}=\left[\frac{2\epsilon_s(V_{CB}+\phi_i)}{qN_{epi}}\right]^{1/2} \tag{7.2.11}$$

集电结空间电荷区宽度与集电极电流的关系可以表示成

$$x_{CB}=\frac{x_{CO}}{(1+J_C/J_1)^{1/2}} \tag{7.2.12}$$

图7.7描述了Kirk效应的影响：当集电极电流增大到接近和超过临界电流时，集电结空间电荷区宽度变窄，中性基区宽度增大。基区的加宽导致电流增益下降，并影响到频率特性。晶体管的频率特性将在下一节讨论。下面继续分析常见的晶体管结构(n型外延层npn晶体管)的Kirk效应。

n型外延层npn晶体管的分析相当复杂，因为集电结空间电荷区的电荷浓度变化很大，p型基区一侧为$-qN_a(x_B)$，n型外延层一侧为$+qN_{epi}$，n^+埋层为$+qN_d(x)$。当集电极电流密度大于$J_1=qN_{epi}v_l$时，外延层的净空间电荷将改变符号，由正变负，导致外延层电场与位置的关系与零电流情形相反。随着集电极电流密度的增大，电力线由小电流时终止于基区受主改变为Kirk效应占主导地位时终止于自由电子；整个集电结空间电荷区分布向重掺杂集电区推

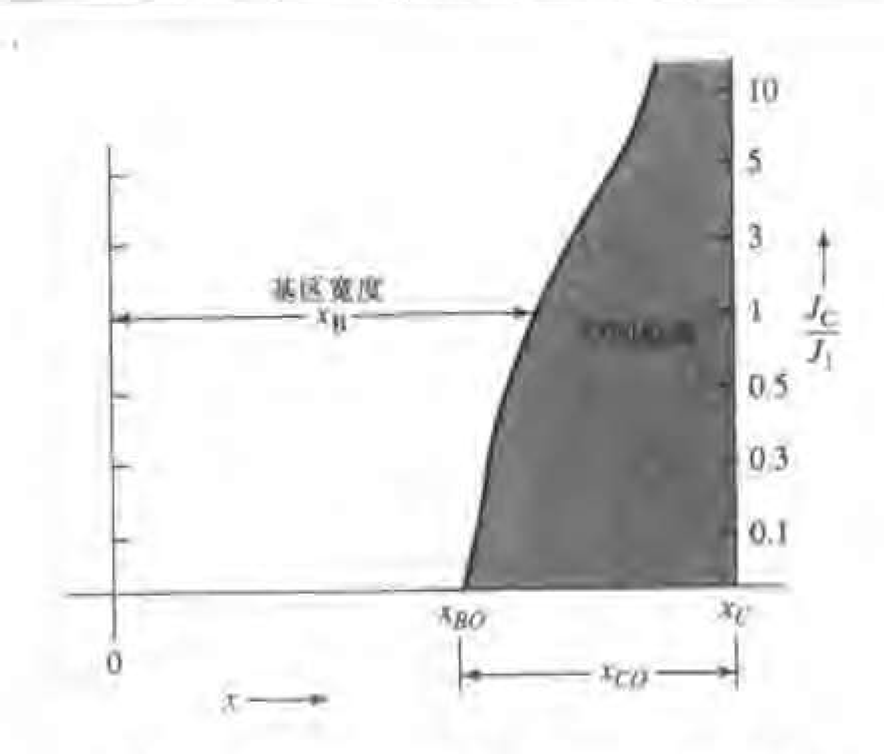


图 7.7　图 7.6 所示的晶体管由于 Kirk 效应引起的基区宽度改变，用等式(7.2.12)计算

进，从小电流时的 pn 结边界推进到高掺杂 n^+ 埋层。在这种大电流情形下，电场随离集电结的距离线性增加，空间电荷区向高掺杂集电区边界压缩（见习题 7.10）。文献[4]对 Kirk 效应进行了计算机模拟，计算结果如图 7.8 所示。

图 7.8(a) 为掺杂分布，图 7.8(b) 和图 7.8(c) 为相应的电场分布和电子浓度分布。小电流时，集电结电场分布近似为三角形（注意：图 7.8(b) 的纵坐标为负电场强度）。该器件的临界电流密度 J_1 为 500A cm^{-2} 数量级。从图 7.8(b) 可看出，在临界电流附近电场强度的极值点向高掺杂的埋层推进。在更高的电流密度下，电场向外延层和埋层之间的边界压缩。图 7.8(c) 的电子浓度分布在发射结附近非常陡峭，以抵消发射结内建电场的作用。随着电流的增加，大电流基区展宽效应非常明显。从该图中还可看出，电流很大时，大部分外延层的电子浓度梯度为常数。

7.2.3　基区电阻

用于放大工作的晶体管的电流增益 β_F 通常很高，可以假定基极电流小到可以忽略，因此也可以忽略基区电阻对晶体管工作的影响。然而，由于 pn 结电流电压的指数关系，基区中很小的电位差都会引起通过发射结的电流密度较大的变化。

考虑如图 7.9 所示的晶体管，当 V_{BE} 从零开始增加时，在基区掺杂浓度 N_B 最低的结处，从发射区注入的电子浓度最高。基极电流沿发射结平面方向横向流动，是多子电流，用于基区复合和向发射区的注入。一般晶体管的基区很薄，典型值不到一微米，因此从基极电极到发射结的电阻是相当大的。流过有源基区不同横截面的基极电流是不同的（见图 7.9），从两侧向中间逐渐减小，因此不能直接利用欧姆定律计算基区电阻。更为重要的是，由于基区电阻上的电压降，导致加在发射结不同位置的偏压不再相同，而是从发射结边缘向中间位置（图 7.9 中 y 轴原点）逐渐减小，相应的发射结电流密度也沿同一方向逐渐减小，形成发射结电流密度的非均匀分布。发射结电流趋向于聚集到发射结边缘附近，这就是电流集边效应。该效应会导致发射结局部发热，降低了器件的最大工作电流，同时也降低了发生大注入效应的临界电流。为降低基区电阻，功率晶体管多采用梳状基极和梳状发射极相互交叉的结构。

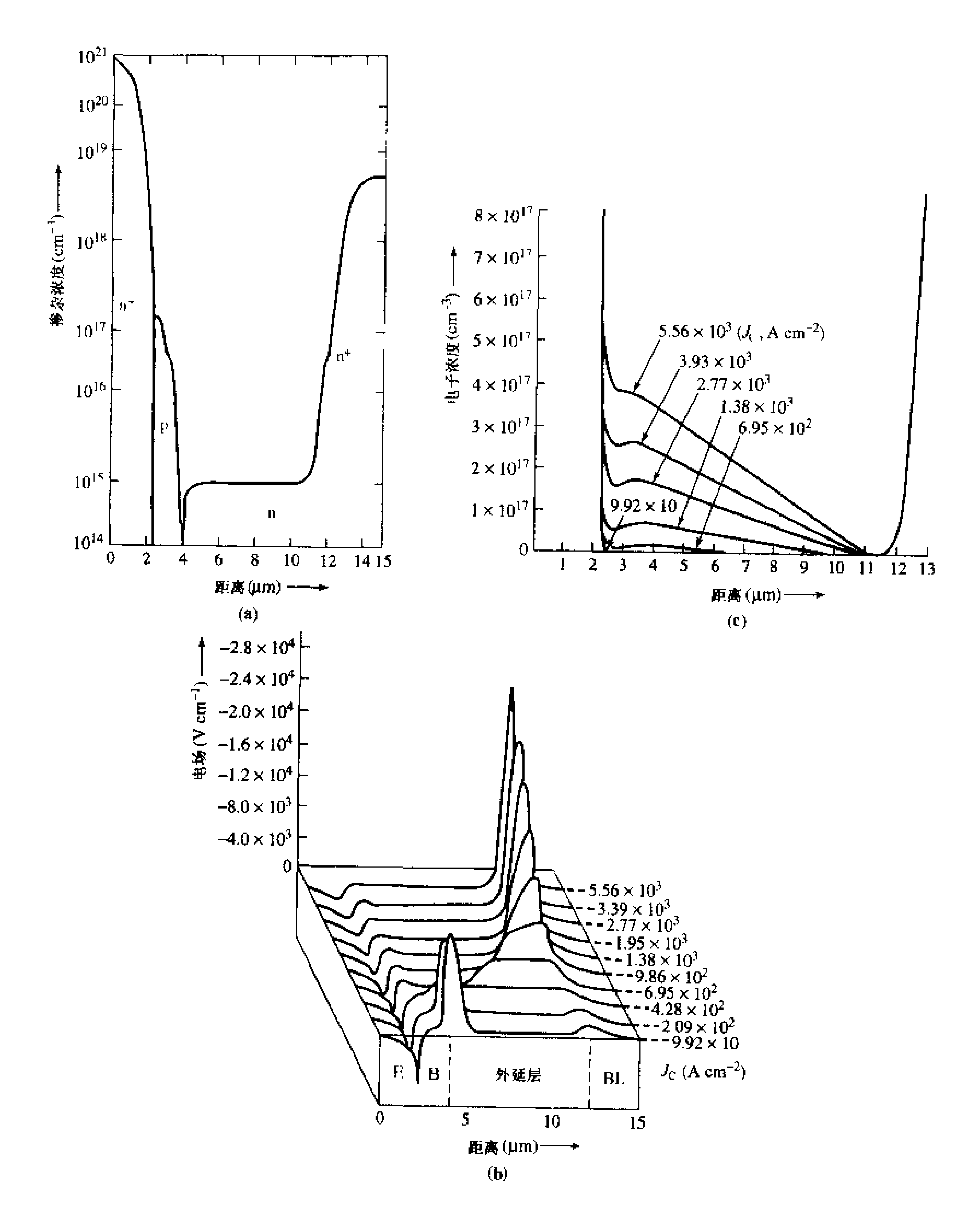


图 7.8 (a)文献[4]中用于计算机模拟 Kirk 效应的 npn 晶体管的掺杂分布;(b)不同的集电极电流密度下的电场分布;(c)不同的集电极电流密度下的基区少子(电子)浓度分布

图 7.10 为集电极电流与基极-发射极电压的关系曲线,其中基区电阻的影响非常明显。由于基区电阻减小了发射结上的电压偏置,我们可以利用等式(7.2.13)来拟合图 7.10 中的

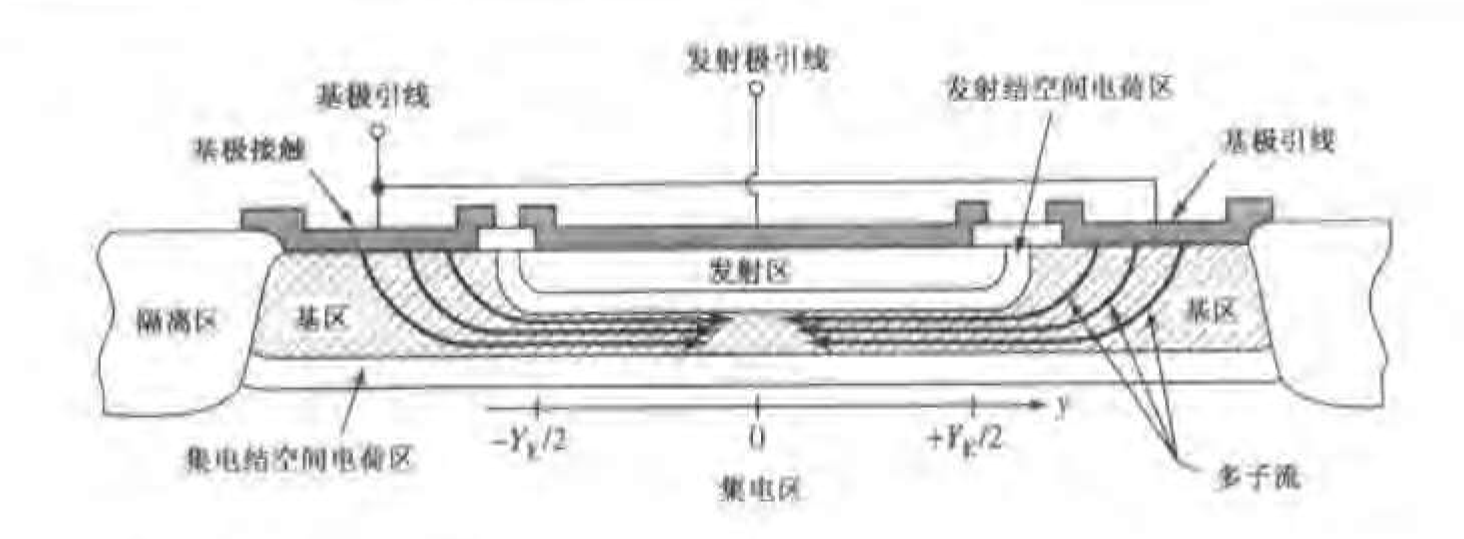


图7.9　工作在正向放大区的晶体管横截面示意图。基极电流从两侧的基极向中间流动，导致发射结不同位置的偏压不同

测量数据

$$I_C = I_S \exp\left[\frac{q(V_{BE} - I_B R_B)}{kT}\right] \tag{7.2.13}$$

由于电流集边，等式(7.2.13)中的基区电阻 R_B（称为基区分布电阻）是可变的。将 I_C、I_B 和 V_{BE} 的测量值代入等式(7.2.13)并计算出 R_B，我们发现随着集电极电流的增大，基区电阻将减小（见图7.11），这是因为电流增大时发射极电流更倾向于聚集到发射结边缘附近，使基极电流实际的输运路径缩短，R_B 下降。

如果只考虑基区分布电阻，那么强集边状态（即所有发射结电流都集中于发射结边缘附近）下 R_B 将趋向于极值，即由无源基区决定的电阻值。如果考虑前面提到的其他大电流效应，R_B 值将更小。

我们可以利用具有空间分布形式的二极管方程来严格求解电流集边效应，但这会使数学推导单调乏味，并且模糊了物理概念。因此，我们采用一种近似的分析方法，将晶体管沿基极电流方向划分成若干子晶体管，每个子晶体管都是忽略了基区电阻的理想晶体管，电流增益与原晶体管相同，满足理想的晶体管方程。每个子晶体管由其分担总饱和电流（等式(7.2.13)中的 I_S）的比例来描述，相邻子晶体管的基极之间通过沿基区的欧姆电阻相串联。增加子晶体管的数目可以提高分析精度，最终可逼近精确解。图7.12将晶体管发射结的左半部分划分成4个子晶体管，然后利用网络理论计算电流电压关系，并利用等式(7.2.13)求出 R_B 和输入电流 I_B 以及输入电压 V_{BE} 之间的关系。基区电阻的进一步分析参考本章习题7.12。

本节分析了晶体管的几种跟偏置有关的物理效应，这些效应限制了放大电路中晶体管的工作范围。通过以上分析我们发现，如果合理地给出基区多子电荷，晶体管的基本关系式6.1.12完全可以描述发射结大注入时的器件特性。但是，如果考虑发射结低偏置、Kirk效应以及基区电阻，第6章的基本理论必须进行修正。下面在讨论晶体管的速度以及模型时将进一步描述这些效应的影响。

为了阐述概念，我们分开讨论了各种大电流效应，认为它们不会同时发生。而实际情况并非如此。晶体管的各种大电流效应之间是相互影响的，其中某些关系可以被预测，例如发射结大注入效应将导致基区电阻减小。但其他效应之间的相互关系非常微妙，若要综合分析需借助计算机辅助分析。文献[6]表明，设计较好的IC晶体管，起主要作用的大电流效应是Kirk

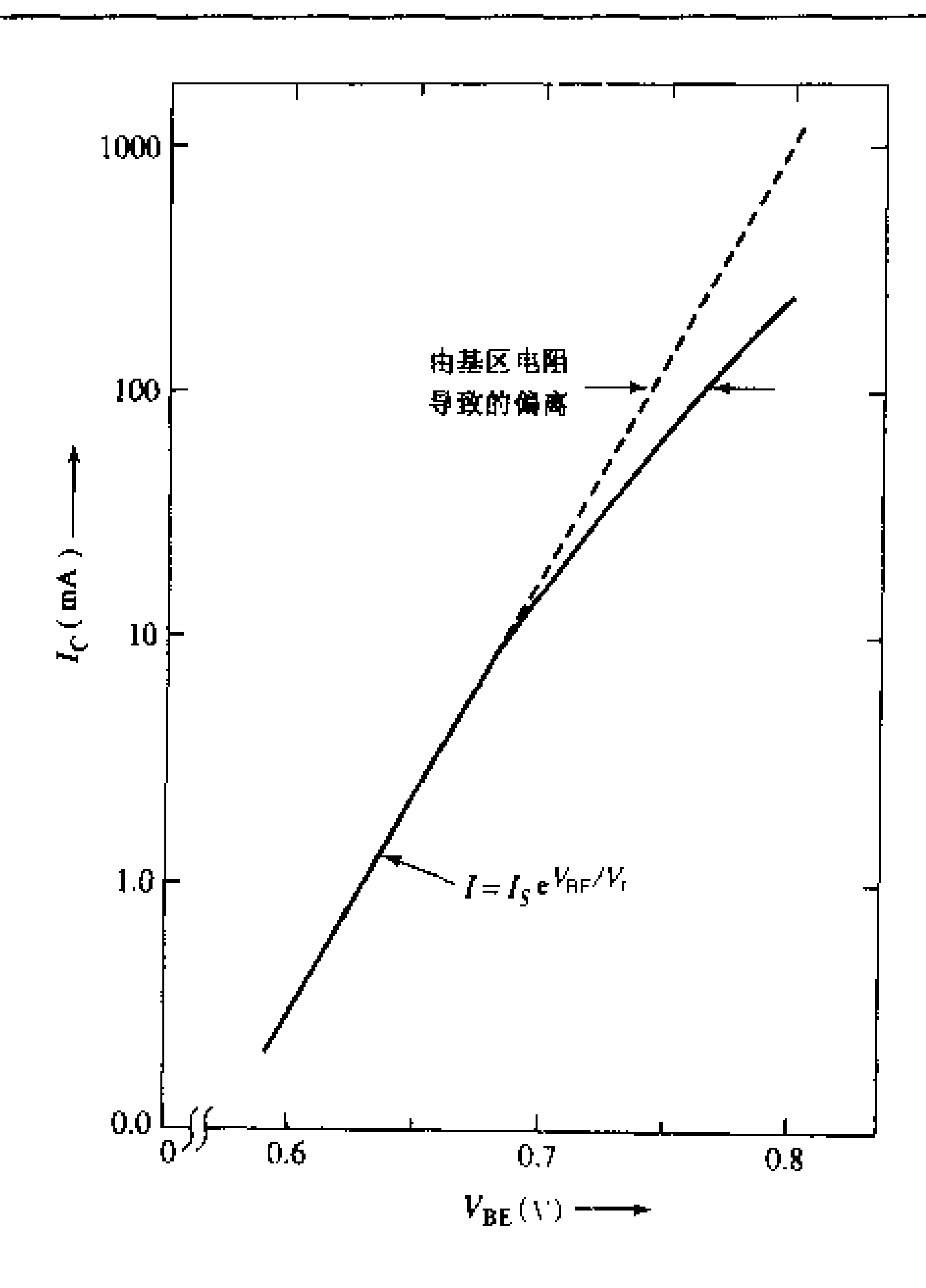


图7.10 晶体管的I_C与V_{BE}的关系曲线,大电流时$I_C \sim V_{BE}$曲线偏离理想特性

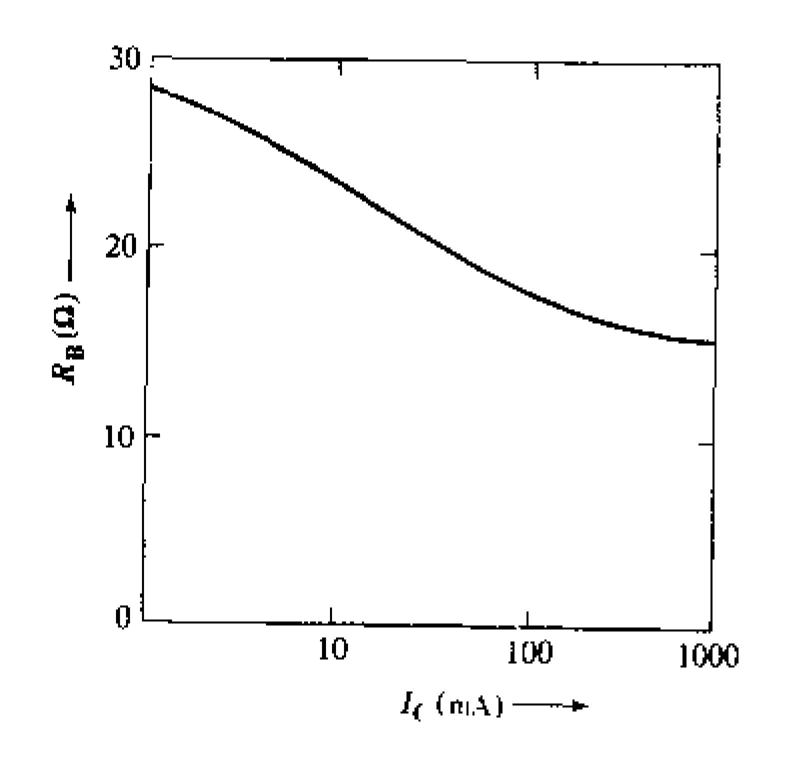


图7.11 对于图7.10的晶体管,基区电阻与集电极电流的关系。基区电阻R_B是利用测量值和等式(7.2.13)计算出来的

效应,即基区展宽效应。

以上分析没有包括晶体管所有的极限效应。除了器件本身的结构会直接影响器件的特性外,IC中的寄生元件也会影响器件性能。在电路分析时通常需要考虑寄生元件导致的极限效

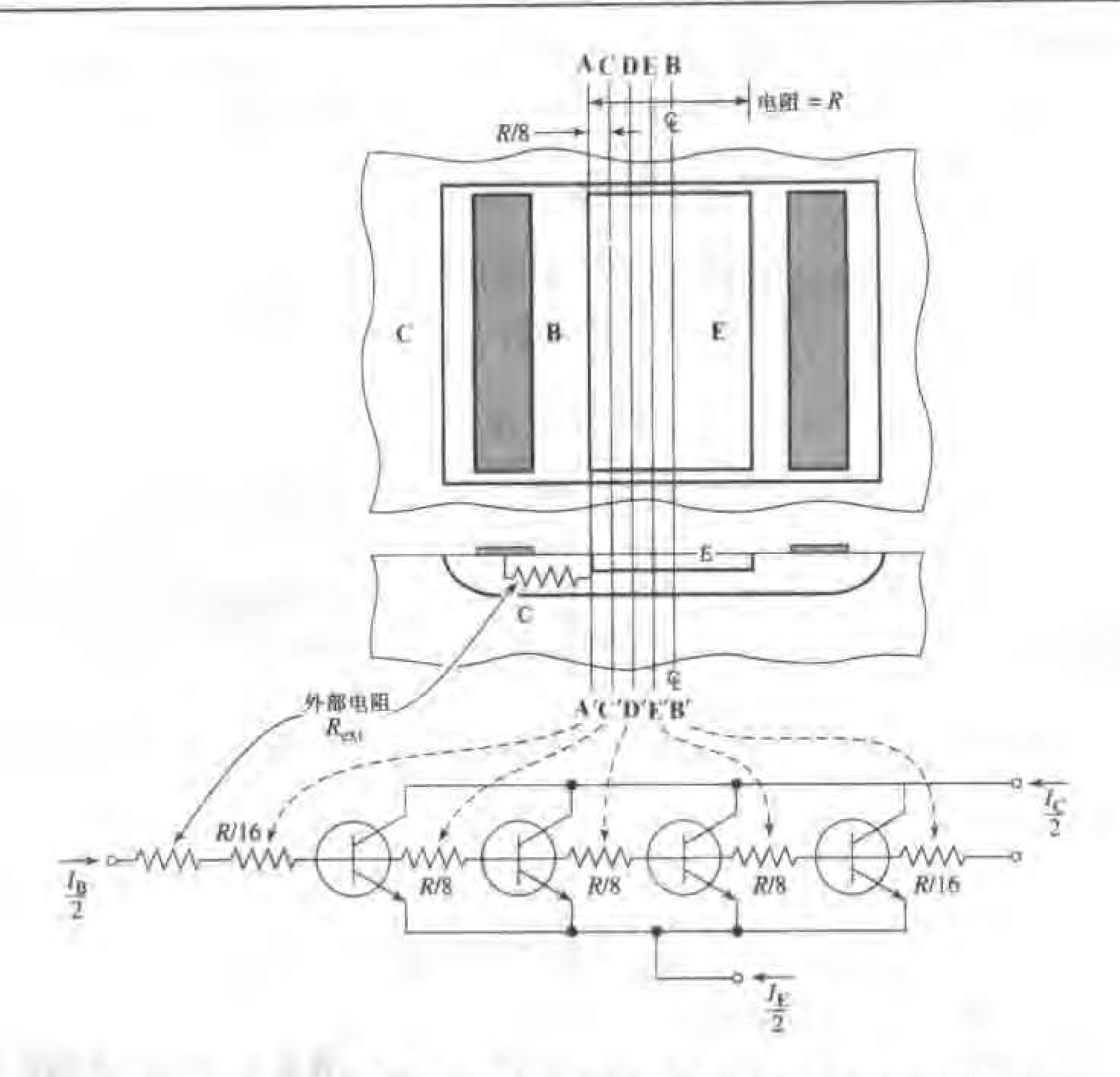


图7.12 基区分布电阻模型。IC晶体管被划分成8个部分（两边各4个），每个部分由理想晶体管和理想欧姆电阻组成[5]。利用网络理论可求出有效基区电阻与基极-发射极偏压之间的关系

应，我们将在7.8节讨论横向pnp晶体管时举例说明。

7.3 基区渡越时间

在正向放大区，少数载流子通过准中性基区的渡越时间是晶体管的一个重要的和基本的参数。与基区渡越时间有关的延时是限制了第一个晶体管性能的主要因素。为降低基区渡越时间所作的研究导致了半导体工艺技术的进步。目前，集成电路双极晶体管的基区渡越时间已经大大降低，以至于基区渡越时间只是影响器件瞬态特性的几个重要的参数之一。然而，考虑基区渡越时间，有助于我们阐述晶体管的设计和性能之间的关系。

我们首先求出放大状态下的晶体管单位面积准中性基区的过剩少子电荷总量 Q_{nB}②

$$Q_{nB}=\int_0^{x_B} qA_E n'(x)\mathrm{d}x \tag{7.3.1}$$

② 二极管n型区的过剩空穴电荷具有相同的定义形式，见等式(5.4.1)。

由于发射结与集电结之间存在晶体管作用,该注入电荷 Q_{nB} 决定了集电极电流的大小(等式(7.1.1))。因此,少数载流子通过基区的基区渡越时间 τ_B 可以表示成 Q_{nB} 与 I_C 的比值

$$\tau_B = \frac{Q_{nB}}{I_C} \tag{7.3.2}$$

图6.1所示原型晶体管的基区渡越时间的计算非常简单,因为 $n'(x)$ 是线性的(等式(6.1.16)),那么 $Q_{nB} = qn'(x)x_B A_E/2$。利用等式(6.1.17)计算出 I_C,那么

$$\tau_B = \frac{x_B^2}{2\tilde{D}_n} \tag{7.3.3}$$

等式(7.3.3)反映了载流子扩散通过基区的时间。与电场作用下的漂移运动相比,载流子的扩散输运速度较低。我们可以这样评估这个问题:利用等式(7.3.3)计算 x_B 为 1μm 的轻掺杂基区($N_a \approx 10^{16}\text{cm}^{-3}$)的 τ_B,得到 $\tau_B \approx 144\text{ps}$;另一方面,当基区存在电场并且电场为常数时,载流子的漂移渡越时间为 $x_B/\mu_n\mathscr{E}$ 或 $x_B^2/\mu_n V$,只要基区的压降超过50mV,漂移时间就会小于扩散时间。

基区杂质任意分布的晶体管,τ_B 的表达式可利用等式(7.1.1)、等式(7.3.1)和等式(7.3.2)得到

$$\tau_B = \frac{\int_0^{x_B} p\,\mathrm{d}x \int_0^{x_B} n'\mathrm{d}x}{\tilde{D}_n n_i^2 \exp(qV_{BE}/kT)} \tag{7.3.4}$$

考虑基区杂质分布对 τ_B 的影响时,等式(7.3.4)非常有用。

将等式(7.3.4)用于大注入状态下的原型晶体管有一定的启发意义。对于放大偏置下的原型晶体管,由于基区载流子浓度随位置线性变化,等式(7.3.4)中的积分非常简单。基区的少子浓度分布为三角形,并且由于基区的准中性,多子浓度分布是从发射结向集电结递降的梯形。图7.13给出了载流子浓度分布图。在这种情形下,利用等式(7.3.4)可得到基区渡越时间为

$$\begin{aligned}\tau_B &= \frac{\left[\frac{1}{2}n'(0)x_B\right]\left[\frac{1}{2}(n'(0) + 2N_a)x_B\right]}{\tilde{D}_n[n'(0)(n'(0) + N_a)]} \\ &= \frac{x_B^2}{4\tilde{D}_n}\left[1 + \frac{N_a}{n'(0) + N_a}\right]\end{aligned} \tag{7.3.5}$$

小注入或中等注入条件下,$n'(0)$ 远小于 N_a,等式(7.3.5)表明 τ_B 趋于 $x_B^2/2\tilde{D}_n$。而大注入条件下 τ_B 趋于 $x_B^2/4\tilde{D}_n$,这说明大注入效应缩短了基区渡越时间。W. M. Webster[7]最早预言了均匀基区晶体管在大注入条件下基区渡越时间将缩短一倍,所以该效应通常称为 Webster 效应。τ_B 减小的物理机构是大注入条件下基区过剩多子电荷形成的电场。为了平衡基区多子由于非均匀分布而导致的扩散,在发射结和集电结之间必须存在电场。该电场平衡了基区多子的扩散趋势,阻止了多子流向集电结;同时,该电场加速了基区少子向集电结的运动,因此缩短了基区渡越时间。均匀基区晶体管的基区少子运动方式在低偏置时只有扩散,而在高偏置时扩散和漂移同时存在。

虽然IC晶体管的基区杂质非均匀分布,但是由7.2节知,大注入时基区将向均匀掺杂的

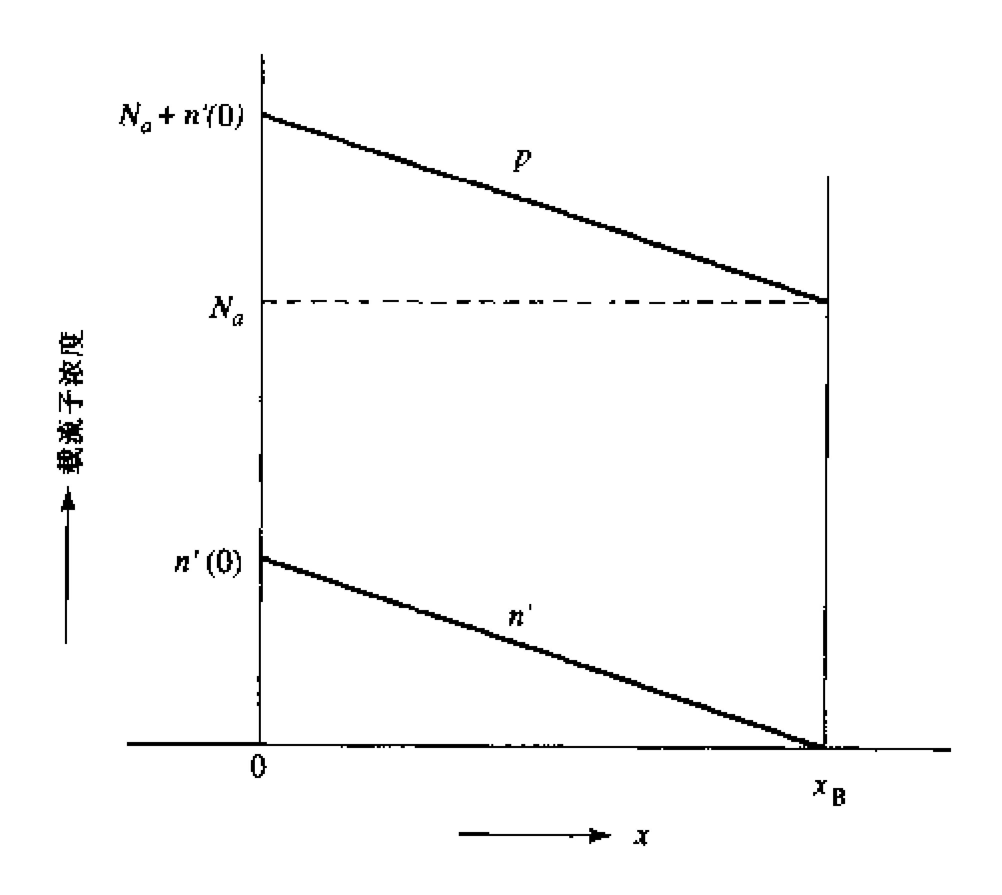


图 7.13 均匀基区晶体管的基区自由载流子浓度分布。基区的电中性导致空穴浓度增加，使空穴浓度大于掺杂浓度

外延层扩展（Kirk 效应），因此，大注入分析时通常取 τ_B 等于 $x_B^2/4\tilde{D}_n$ [3,4]。

τ_B的另一种表达形式可以从等式(6.1.10)和等式(7.3.4)得到。如果取集电结边界的少子浓度$n(x_B)$为零，则从等式(6.1.10)可得到基区少子浓度沿 x 方向的分布

$$n(x) = -\frac{J_n}{q\tilde{D}_n p(x)}\int_x^{x_B} p(\xi)\,\mathrm{d}\xi \qquad (7.3.6)$$

在通常的放大偏置条件下，$n\approx n'$，将等式(7.3.6)代入等式(7.3.4)得到

$$\tau_B = \frac{1}{\tilde{D}_n}\int_0^{x_B}\frac{1}{p(x)}\left[\int_x^{x_B} p(\xi)\mathrm{d}\xi\right]\mathrm{d}x \qquad (7.3.7)$$

将等式(7.3.7)中的参数 x 用 x_B归一化，引入参数 $y=x/x_B$，那么等式(7.3.7)变为

$$\tau_B = \frac{x_B^2}{\tilde{D}_n}\left\{\int_0^1\frac{1}{p(y)}\left[\int_y^1 p(\xi)\mathrm{d}\xi\right]\mathrm{d}y\right\} = \frac{x_B^2}{\nu\tilde{D}_n} \qquad (7.3.8)$$

式中因子 ν 代表了等式(7.3.8)中定积分的倒数，基区非均匀掺杂对基区渡越时间的影响可以由参数 ν 体现。

对于 $p(x)=N_a$的均匀基区晶体管，利用等式(7.3.8)得到 $\nu=2$。当基区杂质为梯度分布时，τ_B可以从 $x_B^2/2\tilde{D}_n$ 减小大约一个数量级。限制 τ_B减小的因素是发射结要维持可以接受的注入效率（要求 $N_a(0)$远低于发射区掺杂浓度），以及 $x=x_B$处要求为非本征的 p 型掺杂（$N_a(x_B)$ 必须大于 N_{depi}）。

7.4 电荷控制模型

等式(7.3.1)引入的用电流和存储电荷的关系来确定时间的概念，不仅可用于计算基区渡越时间，还可扩展到其他方面。电荷控制模型可广泛用于与时间有关的分析[8]。该模型中的控制变量不是电流和电压，而是器件中各区域的控制电荷。本节我们将讨论完整的电荷控制模型，并举例说明其应用。7.6节将利用电荷控制概念来分析限制晶体管频率响应的物理机构。

上一节推导出的等式(7.3.2)是放大偏置晶体管的典型的电荷控制关系。等式 $I_C=Q_{nB}/\tau_B$ 将准中性基区存储的少子电荷与集电极电流 I_C 联系起来，电荷与电流之间为线性关系，比例因子为准中性基区的渡越时间 τ_B。因为等式(7.3.2)只体现了少数载流子通过基区，所以它只是晶体管电荷控制模型的一部分。

放大偏置npn晶体管的工作状态由发射结偏置电压控制。该电压不仅影响 Q_{nB}，还会影响其他的电荷成分。我们主要考虑基区向发射区注入的空穴电荷 Q_{pE} 以及发射结和集电结耗尽区电容的存储电荷 Q_{VE} 和 Q_{VC}。原型晶体管的各电荷成分如图7.14所示。我们首先讨论与稳态基极电流有关的两个注入电荷成分 Q_{nB} 和 Q_{pE}，图7.14中的其他影响BJT动态特性的电荷成分将稍后讨论。由于 Q_{nB} 和 Q_{pE} 都随基极-发射极电压的增加而增加，可用 Q_F 表示 Q_{nB} 与 Q_{pE} 的和(Q_F 随着正向放大偏置电压的增加而增加)。Q_F 的符号为控制电荷(基区多子电荷)的符号，npn晶体管为正，pnp晶体管为负。如果引入特征时间常数 τ_F，就可以根据 Q_F 写出稳态集电极电流表达式(类似等式(7.3.2))

$$I_C=\frac{Q_F}{\tau_F} \tag{7.4.1}$$

需要注意的是，等式(7.4.1)正确的前提是 I_C 必须与 Q_F 呈线形关系(因此 τ_F 为常数)。

Q_F 代表准中性发射区和基区的过剩少子电荷总量，与偏置电压的关系与二极管类似，可以写成

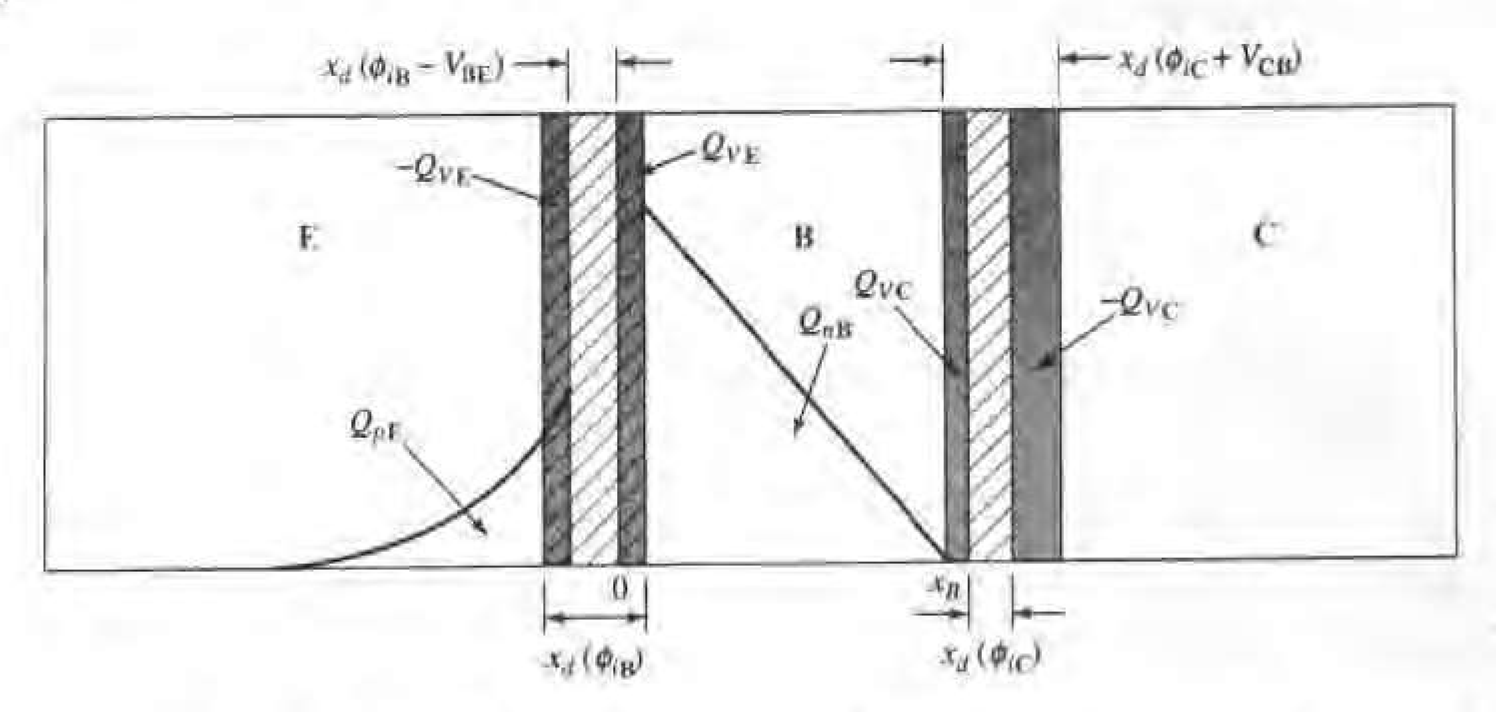


图7.14 原型晶体管电荷控制模型的电荷分布图。电荷 Q_V 代表空间电荷区边界存储的电荷，斜线阴影部分表示热平衡时的空间电荷区

$$Q_F = Q_{FO}\left[\exp\left(\frac{qV_{BE}}{kT}\right) - 1\right] \tag{7.4.2}$$

式中 Q_{FO} 是掺杂分布和器件几何结构的函数。因为 Q_F 大于 Q_{nB}，所以 τ_F 大于基区渡越时间 τ_B（等式(7.3.2)）。在推导完电荷控制模型后，我们将进一步讨论 τ_F。

从基极流入的稳态电流用于准中性基区 Q_{nB} 的复合以及向发射区注入空穴以补充 Q_{pE}，基极电流的大小与二者的和成正比，而这两个成分都与二极管因子 $[\exp(qV_{BE}/kT) - 1]$ 成正比，因此基极电流与 Q_F 成正比（利用等式(7.4.2)）。这样，我们可以写出晶体管基极输入电流的电荷控制表达式

$$I_B = \frac{Q_F}{\tau_{BF}} \tag{7.4.3}$$

从物理结构出发分析发射效率以及发射区和基区的过剩载流子复合要求用更基本的参数来表示 τ_{BF} 和 τ_F，这些内容本书将不涉及。

利用等式(7.4.1)和等式(7.4.3)，可以将稳态电流增益写成两个特征时间的比值

$$\frac{I_C}{I_B} = \beta_F = \frac{\tau_{BF}}{\tau_F} \tag{7.4.4}$$

例如，将等式(7.4.1)、等式(7.4.3)和等式(7.4.4)用于图 6.1 所示的原型晶体管。如果原型晶体管的发射效率非常高，那么 $Q_F \approx Q_{nB}$，并且 $Q_{nB} = \frac{1}{2}qn'(0)x_BA_E$。在这种情形下，因为只有基区复合作用是有效的，所以 $\tau_{BF} = \tau_n$，$\tau_F = x_B^2/2D_n$（利用等式(7.3.3)）。因此，利用等式(7.4.4)可以得到，$\beta_F = 2L_n^2/x_B^2$，其中 $L_n = \sqrt{D_n\tau_n}$。将直流电流增益的这一结论与 6.2 节的分析进行比较：该原型晶体管的 α_F 等于基区输运系数 $[1 - (x_B^2/2L_n^2)]$（等式(6.2.8)），再利用 α_F 与 β_F 的关系 $\beta_F = \alpha_F/(1 - \alpha_F)$，我们发现得到的 β_F 的表达式与用电荷控制分析得到的结论相同。

双极晶体管完整的电荷控制模型需要在电流表达式中增加由于存储电荷随时间变化而出现的其他分量。例如，如果 Q_F 随时间的增加而增加，那么基极电流需要增加一分量 dQ_F/dt。同样，改变发射结和集电结空间电荷区的存储电荷（Q_{VE} 和 Q_{VC}）也需要在基极电流表达式中增加分量。那么，基极电流完整的表达式为

$$i_B = \frac{Q_F}{\tau_{BF}} + \frac{dQ_F}{dt} + \frac{dQ_{VE}}{dt} + \frac{dQ_{VC}}{dt} \tag{7.4.5}$$

等式(7.4.5)的前三项电流从基区流向发射区，最后一项从基区流向集电区。联立等式(7.4.1)和等式(7.4.5)，并利用基尔霍夫电流定律，得到描述放大偏置晶体管的一组电荷控制方程为

$$\begin{aligned} i_C &= \frac{Q_F}{\tau_F} - \frac{dQ_{VC}}{dt} \\ i_B &= \frac{Q_F}{\tau_{BF}} + \frac{dQ_F}{dt} + \frac{dQ_{VE}}{dt} + \frac{dQ_{VC}}{dt} \\ i_E &= -Q_F\left(\frac{1}{\tau_F} + \frac{1}{\tau_{BF}}\right) - \frac{dQ_F}{dt} - \frac{dQ_{VE}}{dt} \end{aligned} \tag{7.4.6}$$

这样，我们推导出一组双极晶体管的电流与电荷关系的线性方程组。这些方程的线性与晶体

管电流电压的非线性关系形成了对比。

图7.15所示的电路图反映了方程组(7.4.6)中的各项。从基极指向发射极的二极管流过稳态电流,其饱和电流为 $I_{FS}=Q_{FO}[(1/\tau_F)+(1/\tau_{BF})]$,正如等式(7.4.2)指出的。描述存储电荷 Q_F、Q_{VF} 和 Q_{VC} 的元件为电容器,穿过电容器的斜线表示这些电容器的电容值随电压改变。现在我们已经有了一组BJT的电荷控制方程组,我们还需要讨论其适用范围和局限性。

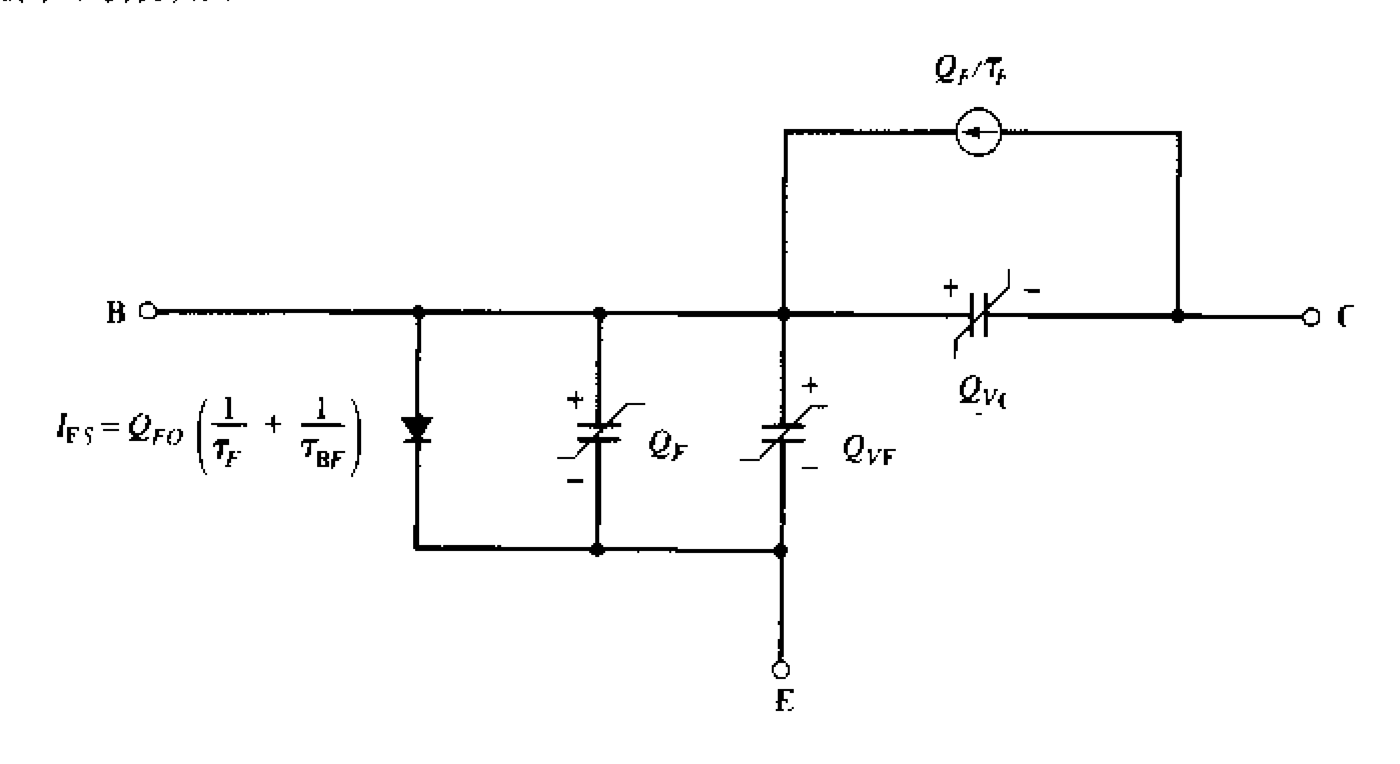


图7.15 放大偏置npn晶体管的电荷控制模型等效电路,该模型考虑了结电荷存储和注入到基区的电荷

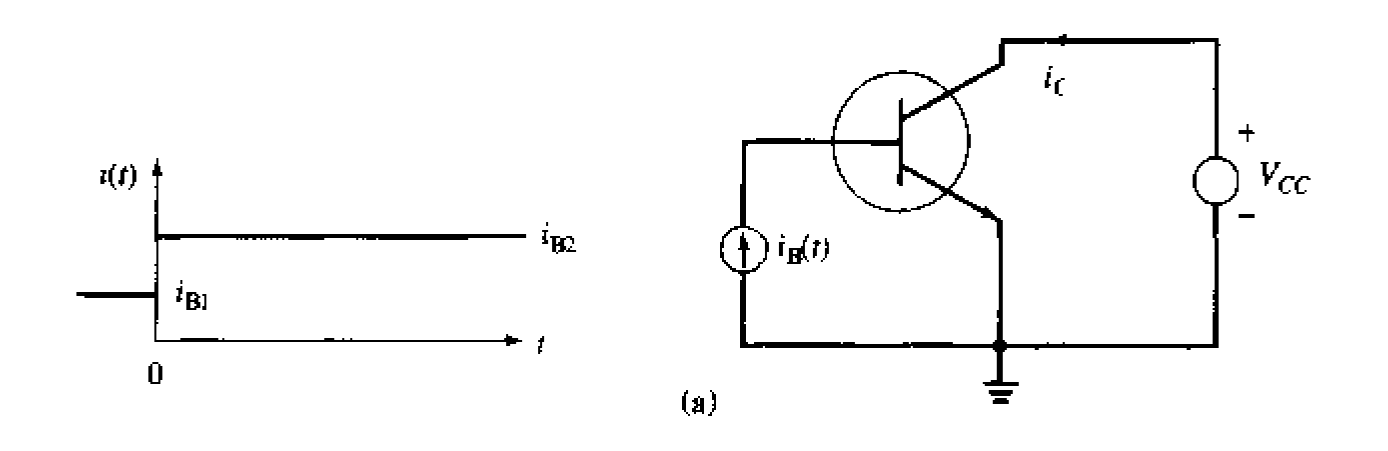
电荷控制分析方法的基本前提是电荷与电流之间存在线性关系,也就是说电荷控制方程中的特征时间与电荷本身以及偏置电压无关。在这个前提下,可以在直流条件下推导特征时间,并将其用于描述动态情形。严格地说,这个前提是不成立的,因为特征时间是电荷的函数。例如,通过发射结注入到基区的少子电荷必须通过基区到达集电结后,才能影响到集电极电流。因此,瞬态情形下集电极电流与基区电荷的关系就与稳态不同。

如果用电荷控制模型分析一个动态问题,求出的电荷的解受到约束,而与"稳态"解不同,这些解有时被称为准静态近似。对于大部分的瞬态分析,电荷控制解通常是相当合理的。当时间尺度达到基区渡越时间的量级,也就是 τ_F 的量级时,电荷控制解将会有相当大的误差。对于大部分的应用来说,时间尺度远大于 τ_F,利用电荷控制模型求解是可以的。下一节我们将举一个特殊的例子来说明电荷控制方法的用途和局限性。

电荷控制模型的应用

在进一步讨论双极晶体管的电荷控制模型之前,我们先举例说明方程组(7.4.6)的应用。我们以图7.16(a)所示的简单电路为例,其电荷控制模型等效电路见图7.16(b)。我们需要分析npn晶体管的基极被一交流电流源驱动时,集电极电流的响应情况。手工计算需要进行适当的简化,首先,可以忽略电流 dQ_{VE}/dt,因为发射结正向偏置时电压的改变很小。这种简化方式通常用于放大偏置的晶体管。另外由于集电极电压为常数,所以电流 dQ_{VC}/dt 也可以忽略。

通过上述简化,基极电流方程只剩下一个未知数,即控制电荷 Q_F



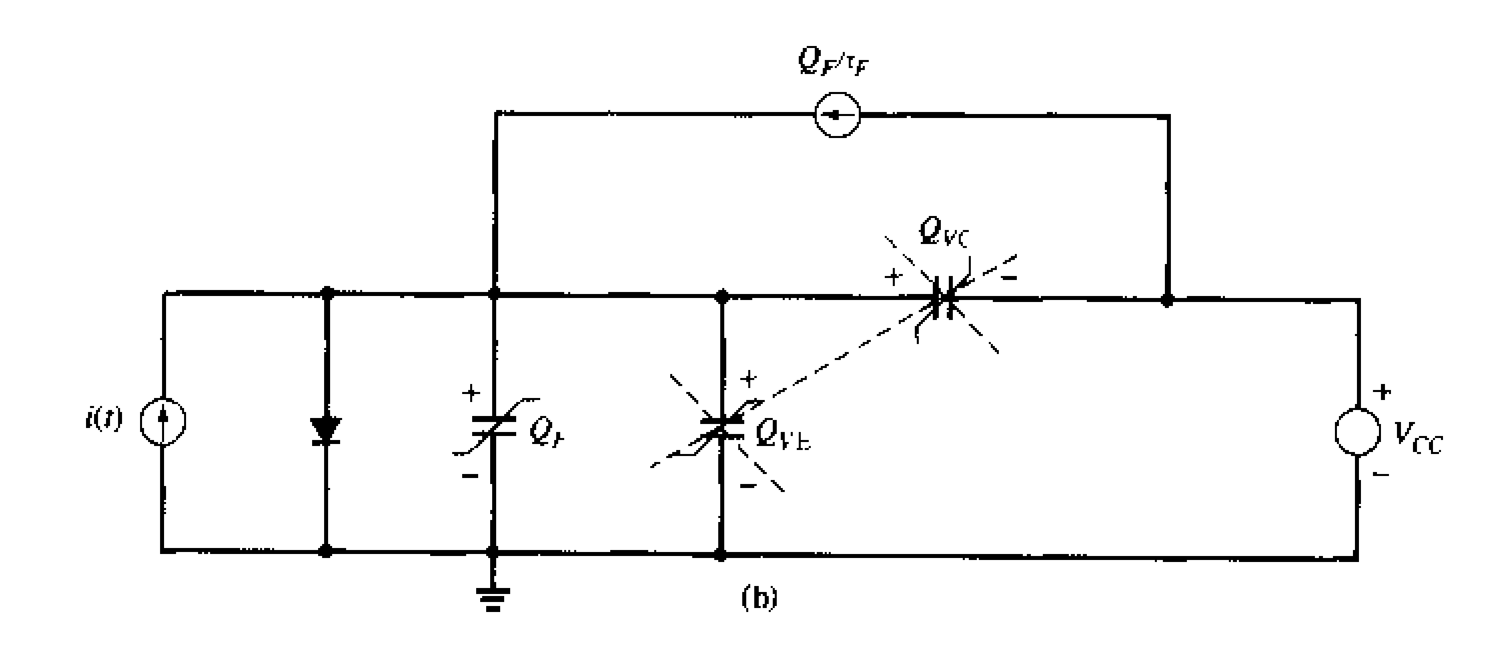


图 7.16 (a)举例说明电荷控制模型的简单电路；(b)等效电路，被删去的元件上流过的电流可以忽略

$$i_B = \frac{Q_F}{\tau_{BF}} + \frac{dQ_F}{dt} \tag{7.4.7}$$

指定输入电流 $i_B(t)$ 为

$$\begin{aligned} i_B &= i_{B1}(t < 0) \\ &= i_{B2}(t > 0) \end{aligned}$$

利用等式(7.4.7)可求出 Q_F的解。当边界条件为 $Q_F(t=0) = i_{B1}\tau_{BF}$和 $Q_F(t\to\infty) = i_{B2}\tau_{BF}$时，$Q_F$的解为

$$Q_F = \tau_{BF}[i_{B2} + (i_{B1} - i_{B2})\exp(-t/\tau_{BF})] \tag{7.4.8}$$

利用前面的近似方法，集电极电流为 Q_F/τ_F，那么集电极电流与时间的关系为(利用等式(7.4.4))

$$i_C = \beta_F[i_{B2} + (i_{B1} - i_{B2})\exp(-t/\tau_{BF})] \tag{7.4.9}$$

这样，集电极电流将以指数关系随时间变化，特征时间常数为 τ_{BF}，如图 7.17 所示。

正如前一节提到的，当时间 t 很小时，等式(7.4.9)给出的瞬态解不准确。例如，在时间零点，等式(7.4.9)预示集电极电流的斜率突变至$(i_{B2} - i_{B1})/\tau_F$，但实际情况是直到发射区注入到基区的电子抵达集电极时，集电极电流才会改变。考虑基区电荷实际分布的更为完善的分析表明，$t=0$ 附近 i_C 不变，基极电流和发射极电流最初的增加量相等。t 等于基区渡越时间 τ_B 时集电极电流开始迅速上升，在 $t=\tau_F$时与电荷控制模型求出的瞬态解达到一致（见图 7.17

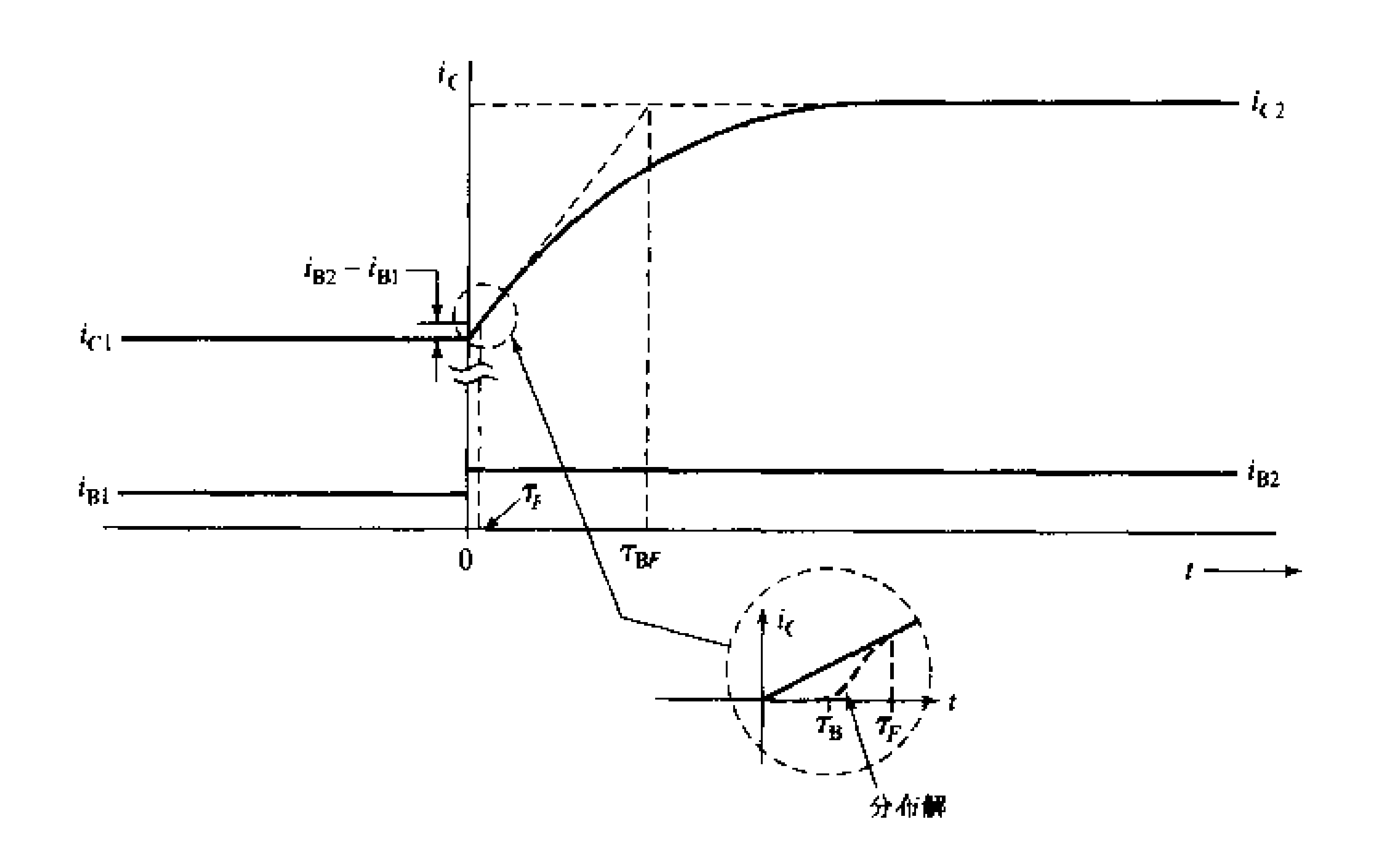


图 7.17 用电荷控制模型求出的图 7.16 所示电路的集电极电流随时间的改变

中的插图)。

这个例子看上去有些不真实,因为我们引入了一些简化,包括电路的选择和一些近似处理。这些物理简化是为了避免复杂的数学处理,但同时也模糊了模型的用处。

如果图 7.16(a)中的集电极通过一个负载电阻 R_L 接到 V_{CC},而不是交流接地,那么问题将变得复杂。由于 V_{CB} 变化,与 Q_{VC} 充电相关的电流不可忽略。如果定义等效电容 C_{jC},其值等于集电极电压变化范围内的 dQ_{VC}/dV_{CB} 的平均值,那么 i_C 很容易求出。如果时间常数 τ_{BF} 用下式替代,那么 i_C 的解仍为等式(7.4.9)(见习题 7.19)

$$\tau'_{BF} = \tau_{BF}\left(1 + \frac{R_L C_{jC}}{\tau_F}\right) \qquad (7.4.10)$$

在推导等式(7.4.10)时,我们注意到集电结电容对基极电流的影响远大于对集电极电流的影响。因此,我们仍然可以假设 $i_C = Q_F/\tau_F$,但是 i_B 必须包括集电结的瞬态电荷存储 dQ_{VC}/dt。

等式(7.4.10)中的第二项表明集电极负载电阻增大了电流的瞬变时间。

大信号模型 电荷控制方程更为广泛的应用是求解大信号瞬态问题,尤其是处理开关问题。开关晶体管通常在关态和饱和区之间转换。因为晶体管所有的工作区都是在两个结上进行电荷的注入和抽取,所以可直接写出大信号开关的电荷控制方程组。注入和抽取过程可以用两组电荷控制方程组的叠加来体现:一组描述正向放大工作(方程组(7.4.6));另一组形式相同,描述反向放大工作。总控制电荷是正向放大偏置电荷 Q_F 和反向放大偏置电荷 Q_R 的叠加。

npn 晶体管完整的电荷控制方程为[3]

3. pnp 晶体管的电荷控制方程组见习题 7.17。

$$
\begin{aligned}
i_E &= -\frac{dQ_F}{dt} - Q_F\left(\frac{1}{\tau_F} + \frac{1}{\tau_{BF}}\right) + \frac{Q_R}{\tau_R} - \frac{dQ_{VE}}{dt} \\
i_C &= \frac{Q_F}{\tau_F} - \frac{dQ_R}{dt} - Q_R\left(\frac{1}{\tau_R} + \frac{1}{\tau_{BR}}\right) - \frac{dQ_{VC}}{dt} \\
i_B &= \frac{dQ_F}{dt} + \frac{Q_F}{\tau_{BF}} + \frac{dQ_R}{dt} + \frac{Q_R}{\tau_{BR}} + \frac{dQ_{VE}}{dt} + \frac{dQ_{VC}}{dt}
\end{aligned}
\tag{7.4.11}
$$

与该方程组相应的等效电路如图7.18所示，电路形式清楚地表明了正向放大与反向放大电荷控制模型的叠加。图7.18中的直流元件与图6.12中的EM模型等效电路一一对应。EM模型的建立方法也是正向放大和反向放大等效电路的叠加。

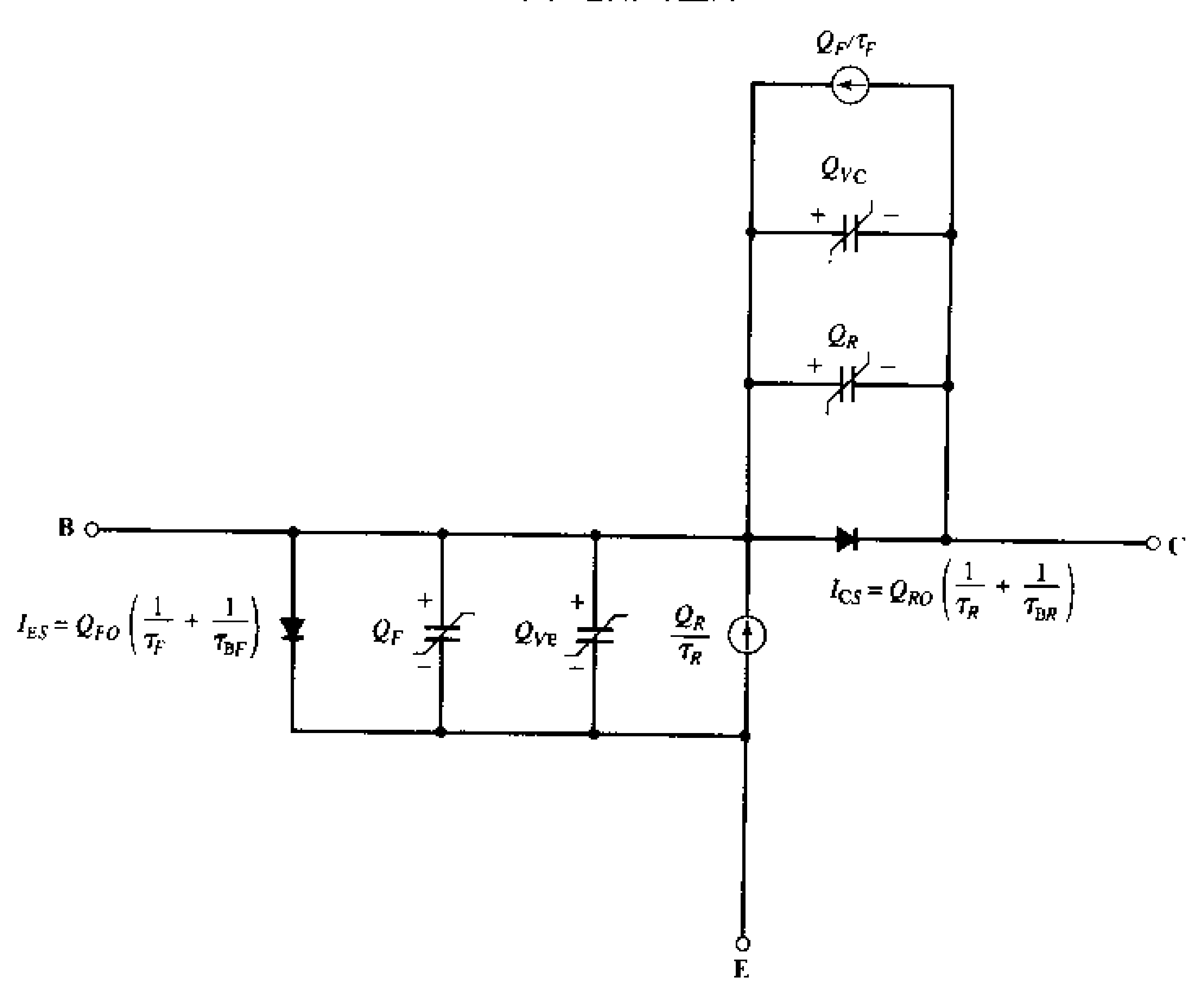


图7.18 完整的双极晶体管大信号电荷控制模型

方程组(7.4.11)中，反向放大偏置相关参数的定义与正向放大偏置类似。例如，与V_{BC}相关的控制电荷Q_R体现了存储在准中性基区和集电区的电荷，其定义为

$$
Q_R = Q_{RO}\left[\exp\left(\frac{qV_{BC}}{kT}\right) - 1\right] \tag{7.4.12}
$$

IC晶体管的集电区掺杂浓度通常与靠近集电结的基区掺杂浓度为同一量级或更低。因此，当晶体管反向放大偏置时，集电区向基区的电子注入效率非常低(为60%到85%)，大量的电荷将存储在集电区。实际上，由于基区很窄，集电区的电荷占Q_R的主要部分。图7.19描述了两种晶体管的饱和区Q_F和Q_R的分布，其中图7.19(a)为集电区轻掺杂的均匀基区晶体管，图7.19(b)为典型的IC晶体管。退饱和时，这两种晶体管都需要抽取大量的存储在集电

区的电荷 Q_C,这将导致相当大的延迟。缩短延迟时间的方法之一是掺入大量的复合中心,通常是在集电区掺金——早期的快速开关集成电路设计常用这种方法提高速度。更好的提高开关速度的方法是不让晶体管进入饱和区,我们可以利用 3.6 节描述的 Schottky 箝位技术来实现。IC 中的 Schottky 箝位晶体管如图 6.17 所示。图 7.20 给出了具有相同尺寸的掺金晶体管和 Schottky 箝位晶体管的开关时间的比较。由于延迟时间的显著缩短,Schottky 箝位晶体管在高速电路中已获得越来越广泛的应用。

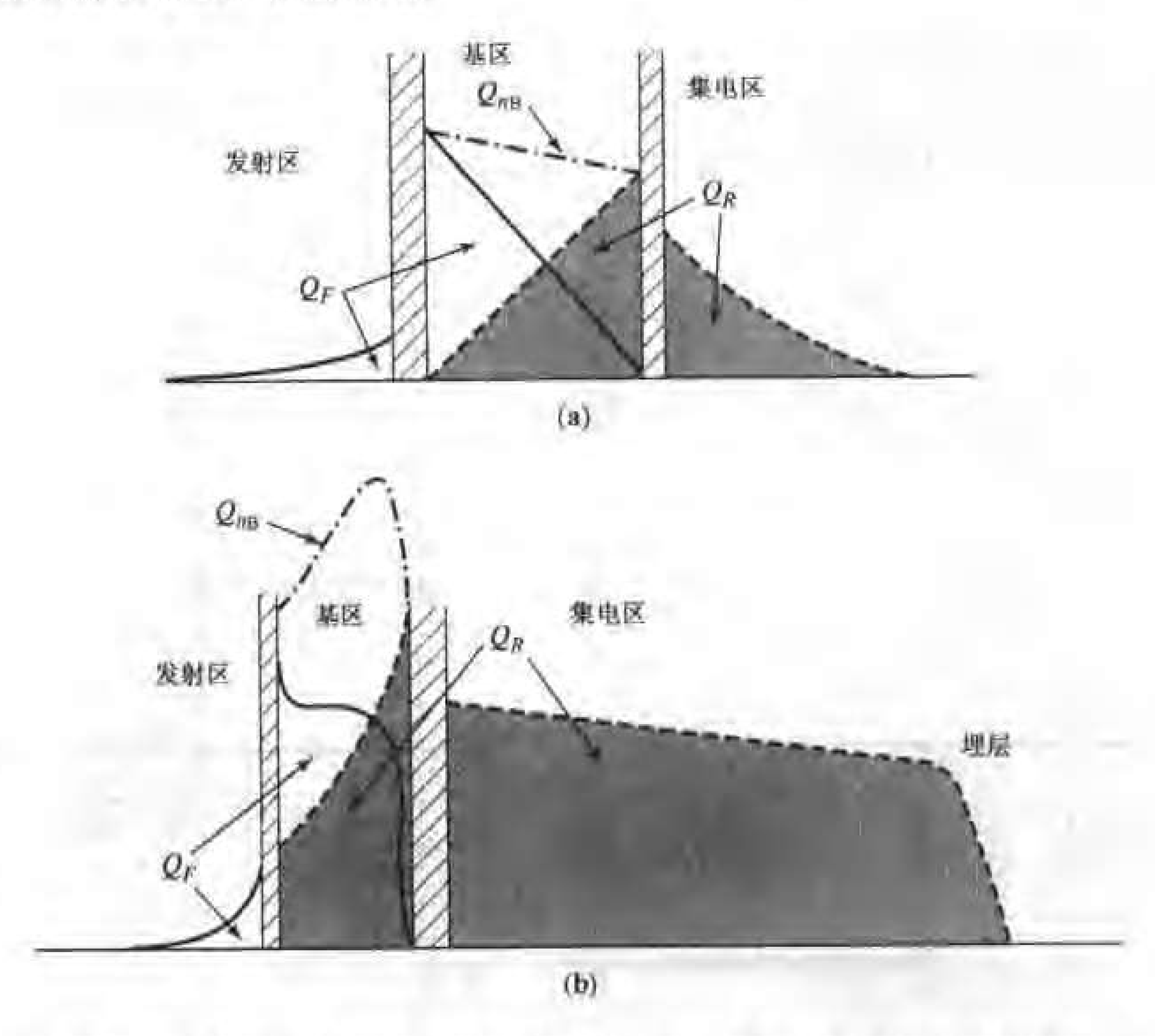


图 7.19 饱和区 Q_F 和 Q_R(虚线)的分布。(a)集电区轻掺杂的均匀基区晶体管;(b)IC 中的外延集电区晶体管。点划线代表基区总电荷 Q_{nB}

瞬态饱和 最后我们利用电荷控制模型求解工作在饱和区的双极晶体管(非 Schottky 箝位晶体管)。这种情形下需要考虑方程组(7.4.11)中的所有项。

分析图 7.21 所示的电路,假设输入电压 $V_S \gg V_{on}$(3.6 节介绍的发射结二极管的"开启电压"),取 V_C 远大于晶体管的饱和压降(等式(6.4.13)),并假设基极驱动足够使晶体管进入饱和工作区,即 $\beta_F V_S/R_S \gg V_C/R_L$。$t=0$ 时刻晶体管开关从关态向开态转换,晶体管进入放大偏置工作区,集电极电流以等式(7.4.9)的形式上升(其中时间常数由等式(7.4.10)给出)。最终的集电极电流不会达到 $\beta_F i_{B2}$[④],因为 V_{BC} 约为 0.5V 时晶体管将进入饱和区,这时集电极电流将被限制在 V_C/R_L,如图 7.22(a)所示。晶体管饱和之前,存储电荷 Q_F 的上升方式与 i_C 类似,如图 7.22(b)所示。尽管晶体管饱和后集电极电流趋于稳定,但电荷 Q_F 和 Q_R 仍随时间改变。直到电荷数值稳定后,晶体管才进入稳定的"开"态。晶体管达到饱和之前,集电结反偏,Q_R 为

④ 译者注:原文误为 $\beta_F i_2$。

零；饱和后 Q_R 开始上升，类似于 Q_F 的上升

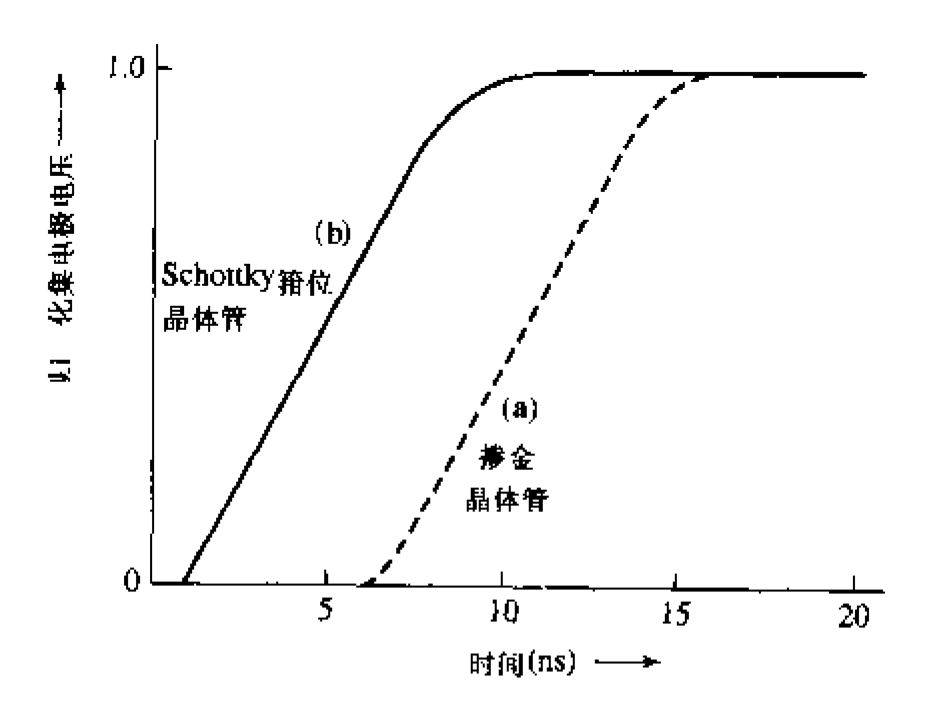


图 7.20　关断时间的比较：(a) 掺金晶体管；(b) Schottky 箝位晶体管。掺金晶体管延迟 7ns，这是因为在改变状态之前需要复合存储电荷[19]

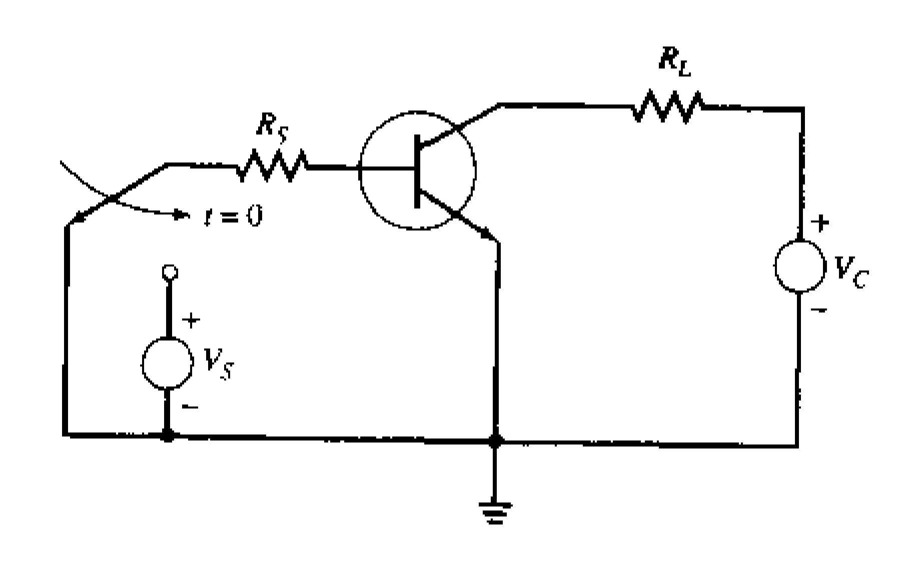


图 7.21　晶体管从关断到饱和的开关电路。输入电压 V_S 远大于发射结"开启电压" V_o，基极驱动电压足够使晶体管进入饱和工作区（$\beta_F V_S/R_S \gg V_C/R_L$）

为分析饱和区特性，需要写出一组相互独立的联立方程组以求解 Q_F 和 Q_R。i_B 和 i_C 的电荷控制方程如下

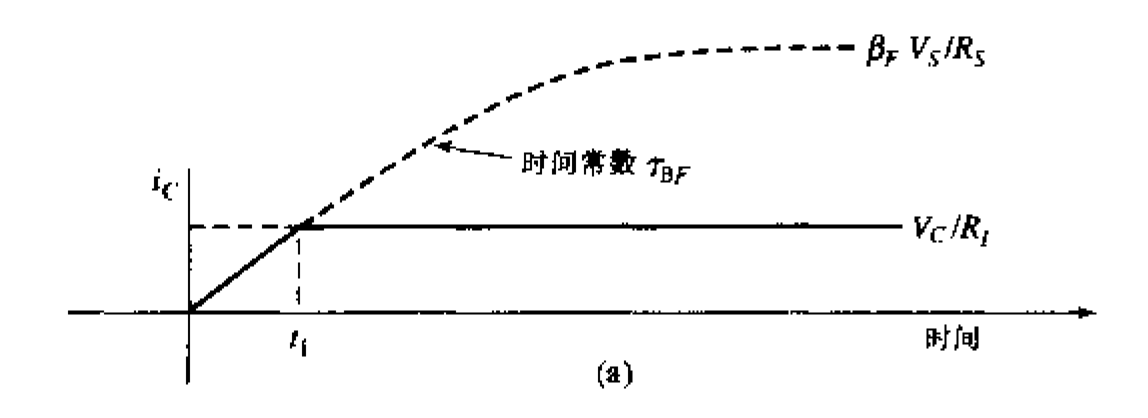
$$
\begin{aligned}
i_B &= \frac{Q_F}{\tau_{BF}} + \frac{dQ_F}{dt} + \frac{Q_R}{\tau_{BR}} + \frac{dQ_R}{dt} \\
i_C &= \frac{Q_F}{\tau_F} - Q_R\left(\frac{1}{\tau_R} + \frac{1}{\tau_{BR}}\right) - \frac{dQ_R}{dt}
\end{aligned}
\tag{7.4.13}
$$

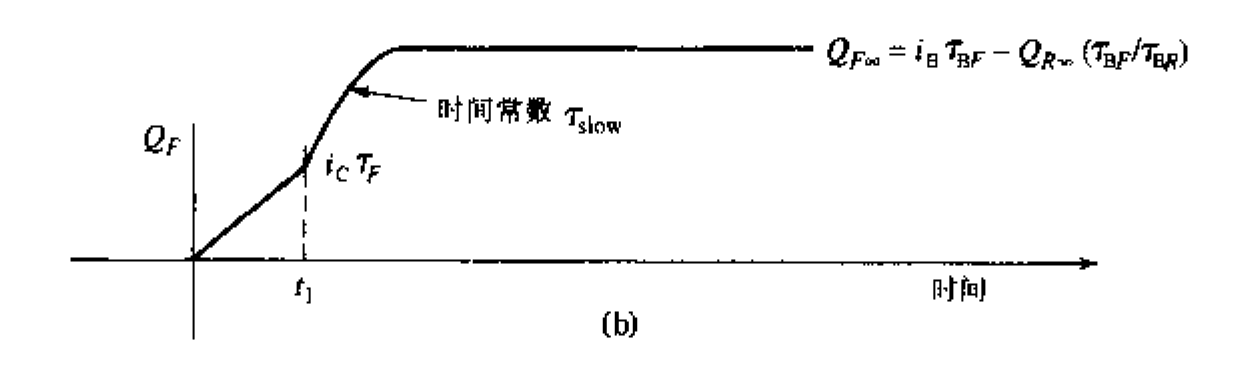
这组方程忽略了与 Q_{VC} 和 Q_{VE} 的充放电相关的电流分量，因为饱和区的 V_{BC} 和 V_{BE} 近似为常数.

微分方程组(7.4.13)的自然频率应满足以下方程

$$
\left(s + \frac{1}{\tau_{BF}}\right)\left(s + \frac{1}{\tau_R} + \frac{1}{\tau_{BR}}\right) + \left(s + \frac{1}{\tau_{BR}}\right)\frac{1}{\tau_F} = 0 \tag{7.4.14}
$$

该二次方程的根为





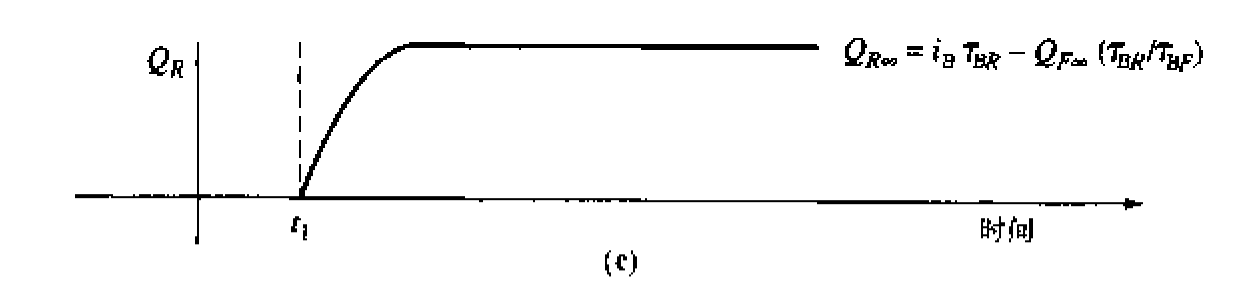


图 7.22 图 7.21 所示的开关电路的瞬态波形:(a) i_C;(b) Q_F;(c) Q_R

$$|s_1| = \frac{1}{\tau_{\mathrm{FAST}}} = \left(\frac{1}{\tau_F} + \frac{1}{\tau_R} + \frac{1}{\tau_{\mathrm{BR}}} + \frac{1}{\tau_{\mathrm{BF}}}\right)$$

和

$$\begin{aligned} |s_2| &= \frac{1}{\tau_{\mathrm{SLOW}}} \\ &= \tau_{\mathrm{FAST}}\left(\frac{1}{\tau_F\tau_{\mathrm{BR}}} + \frac{1}{\tau_R\tau_{\mathrm{BF}}} + \frac{1}{\tau_{\mathrm{BF}}\tau_{\mathrm{BR}}}\right) \end{aligned} \tag{7.4.15}$$

这两个根代表饱和区 Q_F和 Q_R的常规模式的瞬态解,其中根 s_1与晶体管远离到达稳定状态时的瞬态解有关,根 s_2可用于描述决定晶体管饱和状态的电荷 Q_F和 Q_R的瞬态特性。因此,描述饱和区 Q_F和 Q_R的一个较好的近似是(用 Q_{FS}和 Q_{RS}表示)

$$\begin{aligned} Q_{FS} &= (Q_{F\infty} - Q_{F1})\left[1 - \exp\left(\frac{-(t-t_1)}{\tau_{\mathrm{SLOW}}}\right)\right] + Q_{F1} \\ Q_{RS} &= (Q_{R\infty} - Q_{R1})\left[1 - \exp\left(\frac{-(t-t_1)}{\tau_{\mathrm{SLOW}}}\right)\right] + Q_{R1} \end{aligned} \tag{7.4.16}$$

式中 t_1为晶体管开始进入饱和(临界饱和)的时间点,$Q_{F\infty}$和 $Q_{R\infty}$为最终值,Q_{R1}和 Q_{F1}为 $t=t_1$时刻的控制电荷。

饱和区总电荷 $Q_{ST}=(Q_{FS}+Q_{RS})$ 通常被表示成临界饱和态存储电荷 $Q_F=I_C\tau_F$ 与进入饱和区后增加的电荷（超量存储电荷）之和

$$Q_{ST}=I_C\tau_F+Q_S\left[1-\exp\left(\frac{-(t-t_1)}{\tau_{SLOW}}\right)\right] \tag{7.4.17}$$

式中

$$Q_S=Q_{F\infty}+Q_{R\infty}-I_C\tau_F$$

我们可以利用原型晶体管从物理结构上理解这几部分电荷，因为原型晶体管的所有电荷都存储在基区。图 7.23 为该结构的简单示意图，该图要强调的重要特征是，集电极电流稳定时（临界饱和状态），晶体管的开态转换过程并没有结束；只有 Q_F 和 Q_R 达到其最终值时，晶体管才进入稳定状态。

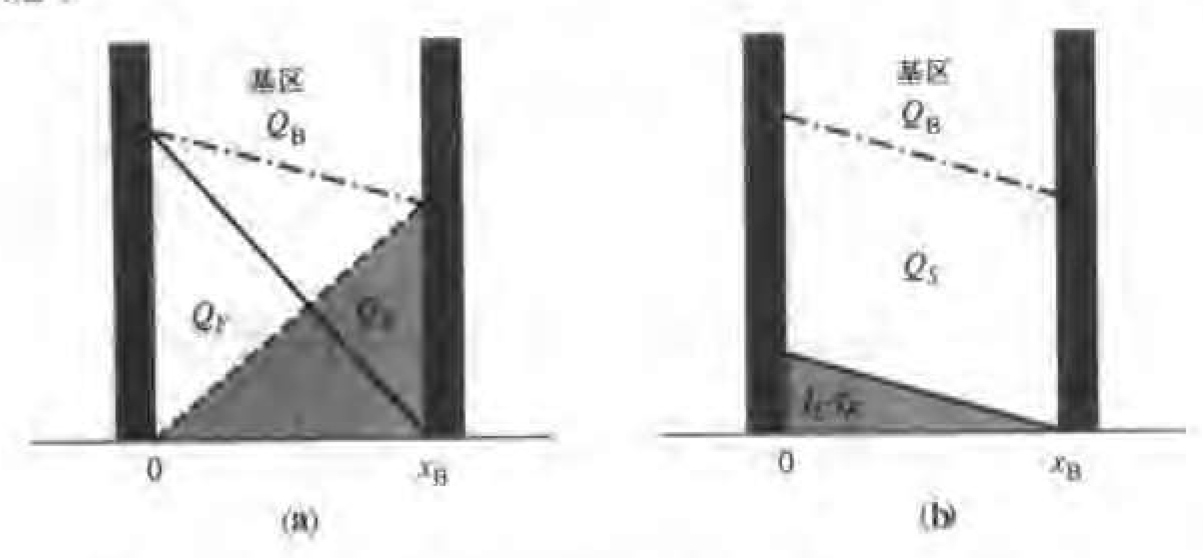


图 7.23　发射区和集电区均为高掺杂的均匀基区晶体管，稳定饱和状态时的基区电荷分布。(a) 控制电荷 Q_F 和 Q_R；(b) 总电荷被分成临界饱和存储电荷 $I_C\tau_F$ 和稳定饱和区超量存储电荷 Q_S

以上分析只考虑了晶体管的开启过程，关断过程的计算与开启过程类似，分为两部分：首先在饱和区求解晶体管从稳定的饱和状态进入临界饱和态（正向放大区的边界），然后在正向放大区求解晶体管从放大状态进入关断状态。

上述讨论有助于表明饱和晶体管电荷控制模型的基本原理，并举例说明了电荷控制模型的应用。由于电荷控制模型提供的一组线性方程组很容易用于计算机辅助分析，因此它的应用非常广泛。参考文献[10]和[11]更为详细地描述了该模型在开关电路中的应用。下一节我们将使用电荷控制模型进行小信号分析。

模型之间的等价　我们已经讨论了双极晶体管的两种模型：6.4 节的 Ember-Moll 模型和本节的电荷控制模型。因为这些模型描述的是相同的器件，我们希望这两种模型的参数之间是相互关联的。例如，在习题 7.20 中我们发现 Ember-Moll 模型参数 α_F 和 α_R 与电荷控制模型参数 τ_F 和 τ_{BF} 的关系分别为

$$\alpha_F=\frac{\tau_{BF}}{\tau_F+\tau_{BF}} \tag{7.4.18}$$

和

$$\alpha_R=\frac{\tau_{BR}}{\tau_R+\tau_{BR}} \tag{7.4.19}$$

Ember-Moll 直流模型中,电荷控制模型的6个基本参数(Q_{FO},Q_{RO},τ_F,τ_{BF},τ_R和τ_{BR})被减少到4个(α_F,α_R,I_{ES}和I_{CS})。

等式(7.4.19)可用于通过测量获得模型参数,也可获得模型的一些特性。例如,利用Ember-Moll模型的互易关系$\alpha_F I_{ES}=\alpha_R I_{CS}$(见等式(6.4.8)),可推导出电荷控制模型的关系式

$$\frac{Q_{FO}}{\tau_F}=\frac{Q_{RO}}{\tau_R} \tag{7.4.20}$$

7.5 晶体管小信号模型

双极晶体管常用于放大小信号,尤其是高频小信号。本节重点讨论晶体管的小信号特性和频率响应特性,并利用电荷控制模型求出晶体管的最高工作频率。

当晶体管工作在正向放大区用于线性放大时,小信号电压通常加在发射结上。如果信号电压远小于热电压$V_t=(kT/q)$,晶体管可以用线性等效电路来描述,这种描述方法被称为晶体管小信号模型,它对放大电路的设计有很大的帮助。利用7.4节的电荷控制关系式很容易推导出小信号模型。

晶体管工作在正向放大区时,集电极电流与基极-发射极电压的关系为

$$I_C=I_S\exp\left(\frac{qV_{BE}}{kT}\right)=I_S\exp\left(\frac{V_{BE}}{V_t}\right) \tag{7.5.1}$$

如果V_{BE}变化,那么I_C的变化为

$$\frac{\partial I_C}{\partial V_{BE}}=\frac{I_S}{V_t}\exp\left(\frac{V_{BE}}{V_t}\right)=\frac{I_C}{V_t}=g_m \tag{7.5.2}$$

式中g_m被称为跨导(定义见等式(4.5.13)),注意g_m与偏置电流成正比。基极电流的变化与发射结偏置电压的关系可直接利用电荷控制关系(等式(7.4.1)和等式(7.4.3))求出

$$\frac{\partial I_B}{\partial V_{BE}}=\frac{\partial(Q_F/\tau_{BF})}{\partial V_{BE}}=\frac{\partial(I_C\tau_F/\tau_{BF})}{\partial V_{BE}}=\frac{\tau_F g_m}{\tau_{BF}}=\frac{g_m}{\beta_F}=\delta g_m \tag{7.5.3}$$

(τ_F和τ_{BF}的比值δ称为缺损因子。从等式(7.4.4)可得出$\delta=\beta_F^{-1}$)

基区少子电荷随基极-发射极电压的变化为

$$\frac{\partial Q_F}{\partial V_{BE}}=\frac{\partial(I_C\tau_F)}{\partial V_{BE}}=g_m\tau_F=C_D \tag{7.5.4}$$

式中符号C_D(通常称为扩散电容)代表由于注入少子电荷的改变引起的电容效应。

如果等式(7.5.2)、等式(7.5.3)和等式(7.5.4)中的电压和电流的增量代表交流信号,那么这一系列等式可以用图7.24所示的等效电路表示。小信号电流和电压增量用小写字母表示。发射结输入回路是一个并联的RC网络,其时间常数为$\beta_F\tau_F$,也就是基区时间常数τ_{BF}。

集电极输出回路是一个受输入电压控制的电流源。在给定的v_{BE}下,输出电流与g_m成正比,因此也依赖于直流工作点I_C(等式(7.5.2)[5])。图7.24所示的一级近似等效电路中,输入不

5 译者注:原文误为7.5.1。

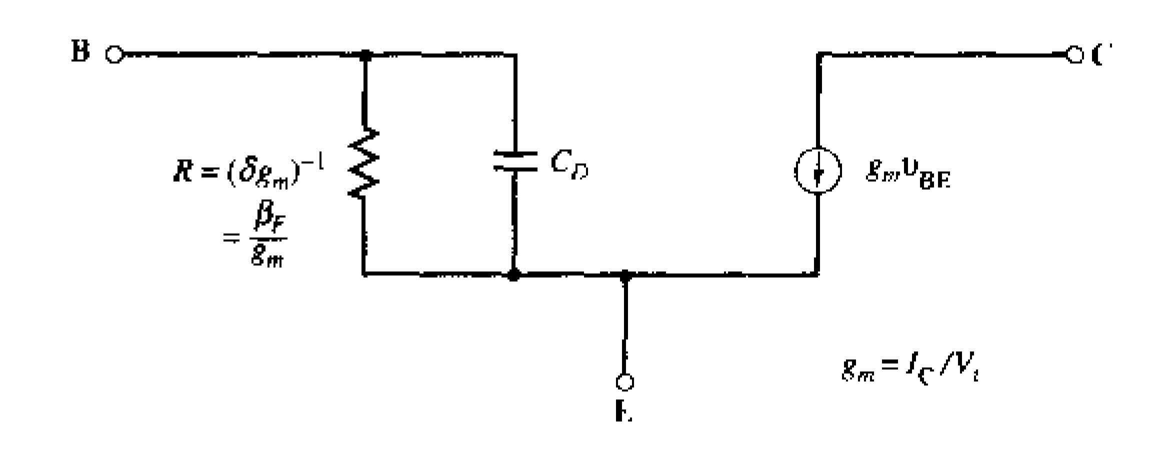


图7.24 双极晶体管正向放大工作时的小信号等效电路,只包括一级近似

受输出影响,并且输出不受集电极-基极电压的影响。由7.1节知,由于Early效应,集电极-基极电压的大小对集电极电流是有影响的。I_C与V_{CB}的关系可用Early电压V_A描述。根据小信号参数的定义,有

$$\left|\frac{\partial I_C}{\partial V_{CB}}\right| = \frac{I_C}{|V_A|} = \frac{g_m V_t}{|V_A|} = \eta g_m \tag{7.5.5}$$

式中新引入了参数$\eta \equiv V_t/|V_A|$,代表V_{CB}改变引起I_C的变化与V_{BE}改变引起I_C的变化之间的比值

I_C随V_{CB}变化还会引起控制电荷Q_F随V_{CB}变化。利用电荷控制关系式(7.4.1)得出

$$\left|\frac{\partial Q_F}{\partial V_{CB}}\right| = \left|\frac{\partial (I_C \tau_F)}{\partial V_{CB}}\right| = \tau_F \eta g_m = \eta C_D \tag{7.5.6}$$

基区少子电荷的任何改变都会导致基极和集电极电流的变化。基极电流随V_{CB}的改变为

$$\left|\frac{\partial I_B}{\partial V_{CB}}\right| = \left|\frac{\partial (Q_F/\tau_{BF})}{\partial V_{CB}}\right| = \frac{\eta g_m \tau_F}{\tau_{BF}} = \eta \delta g_m \tag{7.5.7}$$

等式(7.5.5)、等式(7.5.6)和等式(7.5.7)中的关系可以通过在等效电路图7.24中加入三个元件来体现,如图7.25所示。等式(7.5.5)计算的是集电极-基极电压的变化对集电极到发射极的电流的影响,因此用一个集电极-基极电压控制的电流源等效。该电流源从集电极流向发射极,因为由7.1节知,V_{CB}增加将导致I_C增加。等式(7.5.6)计算的是集电极-基极电压对存储电荷的影响,可以用一个连接在集电极和基极之间的电容等效。等式(7.5.7)计算的是集电极-基极电压的改变引起的发射结电流的变化,可用发射极指向基极的电流源等效。该方向将导致基极电流减小,这是因为Q_F减小了。

电路分析时,图7.25中的等效电路可用两种方式进行简化。首先,包含v_{CB}的电流源可以用图7.26所示的方式等价为连接在集电极和基极之间的电流源,这样做之后,有的电流源的激励电压就是其两端的电压,因此可以用无源器件替代它。第二种方式是利用恒等式$v_{CE} = v_{CB} + v_{BE}$重新表示受电压控制的有源器件。完成这些操作并合并并联电阻,等效电路可简化成图7.27所示的形式,这种小信号等效电路通常被称为混合Π等效电路。混合是因为有源器件是电压控制的电流源,联系了不同量纲的参量,Π是因为等效电路形式类似大写希腊字母Π

为了更精确地描述晶体管的小信号特性,还需要考虑两个效应。首先是7.2节讨论过的基区电阻。由于电流集边效应,直流基区电阻R_B是集电极电流的函数(等式(7.2.13)),典型的npn晶体管的基区电阻R_B和I_C的关系如图7.11所示。为了在等效电路中反映该效应,必

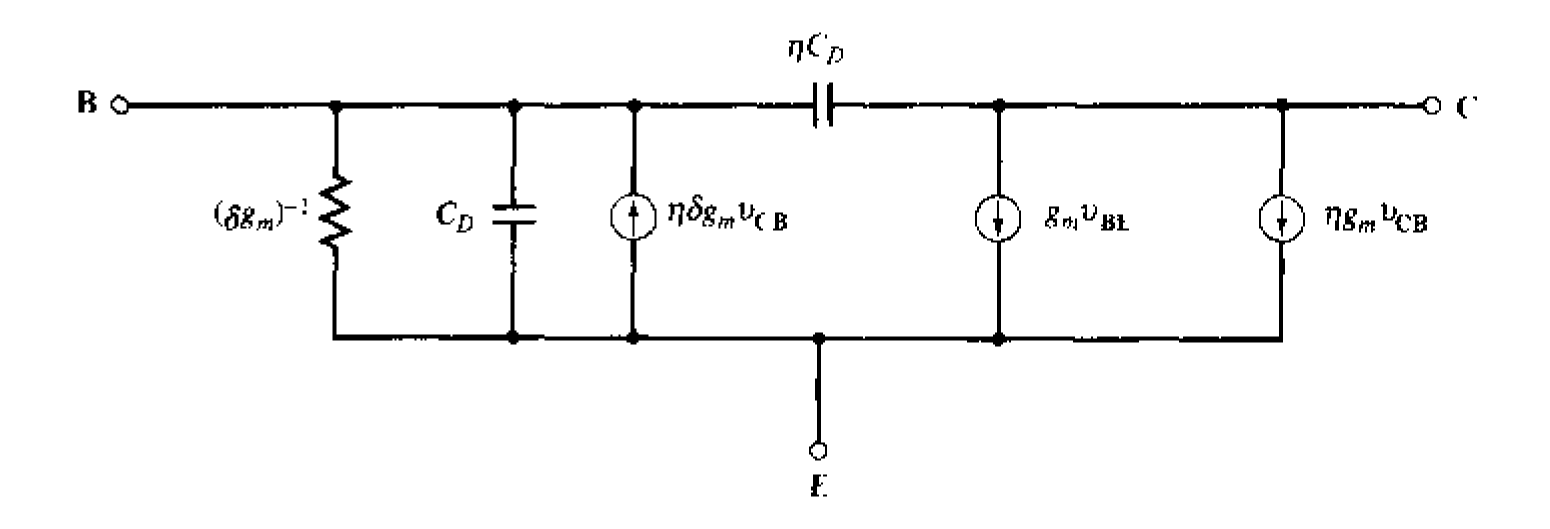


图 7.25 双极晶体管小信号等效电路，考虑了 Early 效应

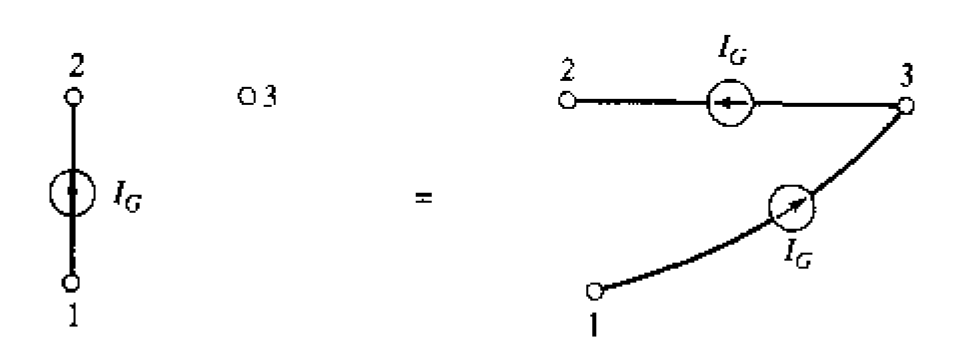


图 7.26 描绘两节点间电流的两个等价的电路

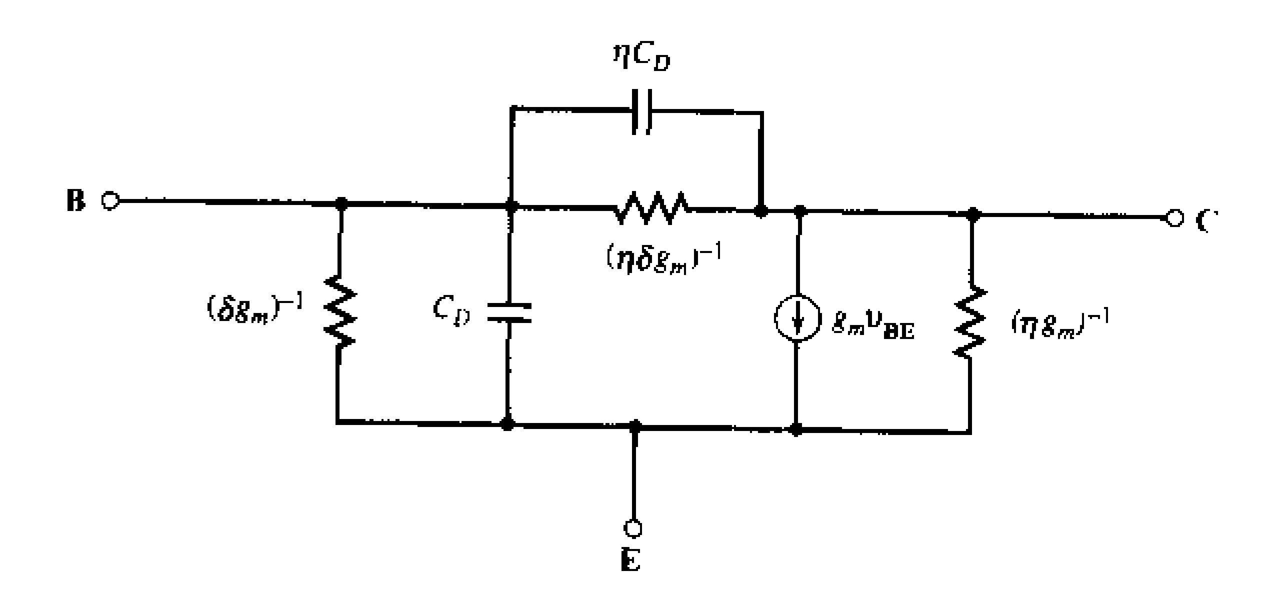


图 7.27 简化后的小信号等效电路，考虑了 Early 效应．通常 $\delta \ll 1, \eta \ll 1$

须考虑小信号电流电压关系，即分析 V_{BE}、I_B、I_C 和 R_B 之间的函数关系。V_{BE} 对 I_C 的全微分包括三项

$$\frac{\mathrm{d}V_{BE}}{\mathrm{d}I_C} = \left.\frac{\partial V_{BE}}{\partial I_C}\right|_{I_B,R_B} + \left.\frac{\partial V_{BE}}{\partial I_B}\right|_{I_C,R_B}\left(\frac{\mathrm{d}I_B}{\mathrm{d}I_C}\right) + \left.\frac{\partial V_{BE}}{\partial R_B}\right|_{I_C,I_B}\left(\frac{\mathrm{d}R_B}{\mathrm{d}I_C}\right) \tag{7.5.8}$$

从等式(7.2.13)得到

$$V_{BE} = I_B R_B + V_t \ln\left(\frac{I_C}{I_S}\right) \tag{7.5.9}$$

代入等式(7.5.8)，得到

$$\frac{\mathrm{d}V_{BE}}{\mathrm{d}I_C} = \frac{V_t}{I_C} + R_B\frac{\mathrm{d}I_B}{\mathrm{d}I_C} + I_B\frac{\mathrm{d}R_B}{\mathrm{d}I_C} = \frac{1}{g_m} + \delta R_B + I_B\frac{\mathrm{d}R_B}{\mathrm{d}I_C} \tag{7.5.10}$$

式中利用了g_m和δ的定义，同时保留了R_B的微分项。在实践中，该微分项可以从类似图7.11的曲线中获得。

为了将等式(7.5.10)体现在等效电路中，我们来求解基极输入电阻R_I

$$R_I = \frac{dV_{BE}}{dI_B} = \frac{dV_{BE}}{dI_C} \cdot \frac{\partial I_C}{\partial I_B} = \frac{dV_{BE}}{dI_C} \cdot \frac{1}{\delta} \tag{7.5.11}$$

将等式(7.5.10)代入等式(7.5.11)，得到

$$R_I = \frac{1}{\delta g_m} + \left(R_B + \frac{I_B}{\delta}\frac{dR_B}{dI_C}\right) \tag{7.5.12}$$

因此，考虑基区电阻，必须将以下电阻与基极-发射极电阻$(\delta g_m)^{-1}$串联

$$r_b = R_B + I_C \frac{dR_B}{dI_C} \tag{7.5.13}$$

为了集中考虑基区电阻的作用，我们将图7.27中的等效电路简化成如图7.28所示的形式，其中忽略了电容（低频时无效）和大电阻。显然，如果考虑基区电阻，输出回路的电流源不再由基极-发射极电压驱动，而是由内部节点电压驱动。图7.28所示的等效电路在加入了基区电阻后，电流增益是否能合理建模，这个问题作为习题留下，请读者推导。

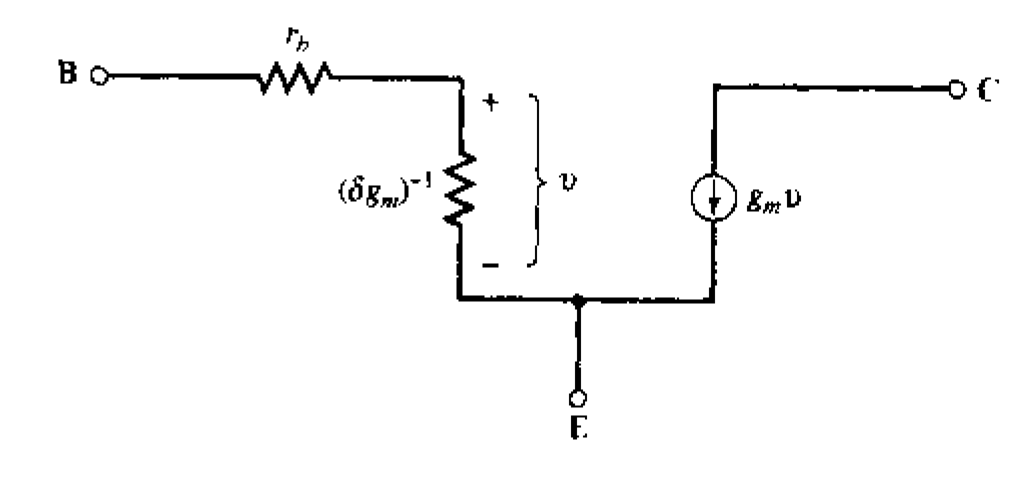


图7.28 考虑了基区电阻的低频小信号等效电路

目前我们只考虑了直流和低频情形下的基区电阻。高频时由于基区电阻和结电容类似分布传输线，问题将更为复杂。用并联的RC网络替代r_b来模拟该效应将更为精确，详细分析见参考文献[5]。

最后需要加入到混合Π等效电路中的元件是势垒电容，这些电容与图7.27中的发射结和集电结电容并联，通常用符号C_{jE}和C_{jC}表示。势垒电容值可以利用第4章的势垒电容公式以及7.4节的电荷控制模型进行计算。发射结势垒电容为$C_{jE} = dQ_{VE}/dV_{EB}$，集电结势垒电容为$C_{jC} = dQ_{VC}/dV_{CB}$。为提高精度，可以将集电结势垒电容分为两部分：非本征晶体管对应的跨接到基极的电容（称为C_{xC}）；连接到本征晶体管基极节点（连接δg_m的节点）的电容C_{jC}。这样划分是因为其中一部分集电结势垒电容并不通过基区电阻充放电。包含这些效应的完整的等效电路如图7.29所示。

这个电路非常复杂，幸运的是我们很少需要手工计算完整的电路。我们可以在指定条件下忽略某些元件以简化电路，或者利用计算机辅助分析。

频率响应 本节通过对一些物理效应的考虑，得出了放大偏置晶体管的小信号混合Π等

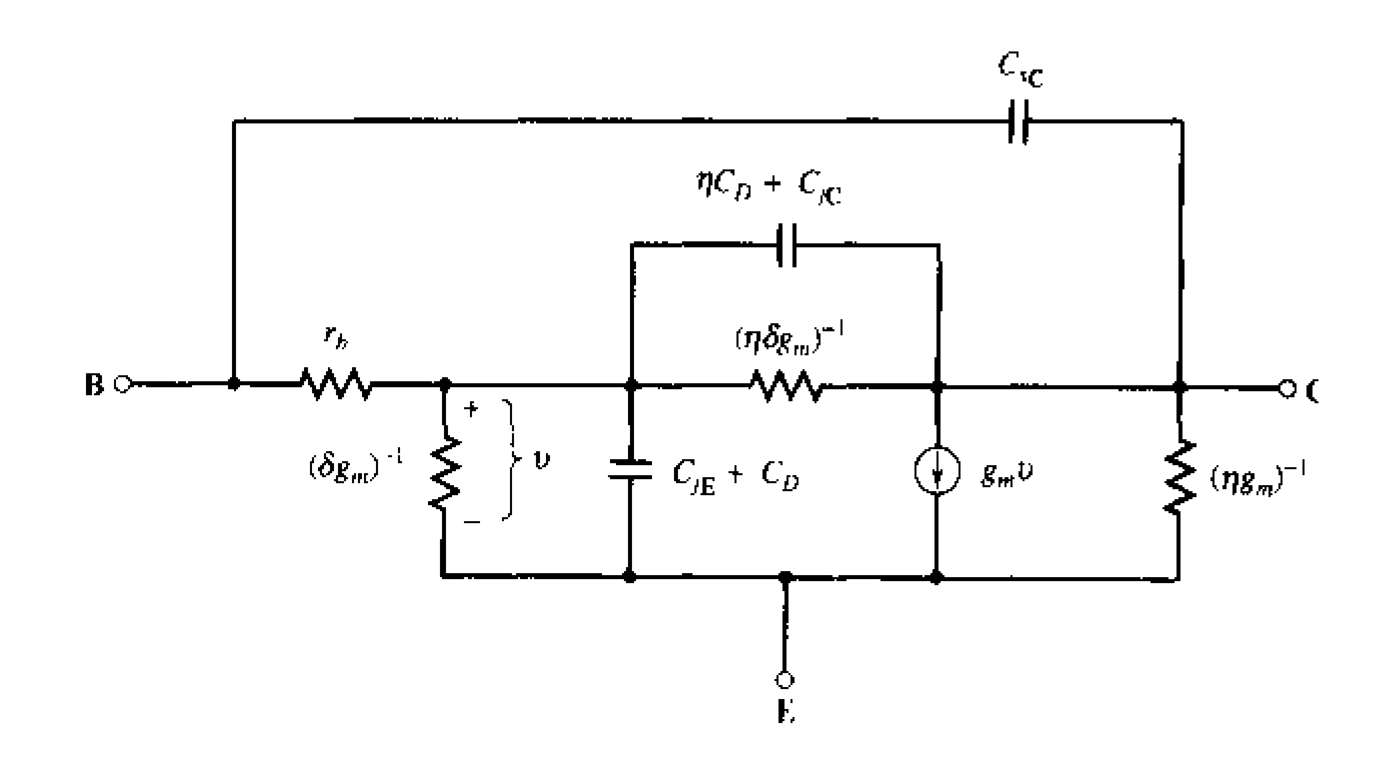


图 7.29 考虑了势垒电容和基区电阻的混合 Π 等效电路

效电路。我们可以用这个电路来研究晶体管的频率响应,尤其是最高工作频率。

将小信号混合 Π 等效电路与电荷控制模型相结合,通过测量短路电流增益与频率的关系,可以确定电荷控制参数 τ_F。下一节我们将扩展电荷控制模型,考虑其他电荷在晶体管中的运动以及由此引起的附加延迟 这些考虑将提供一个更为复杂的,并能从物理上反映晶体管频率限制的参量。

如果集电极只加载直流偏置,那么小信号等效电路中的集电极交流接地。在这种条件下,可忽略图 7.29 的电路中反映 Early 效应的元件,另外还可忽略影响不大的基区电阻。简化后的电路如图 7.30 所示,该电路的电流增益 i_C/i_B 为

$$\frac{i_C}{i_B} = \frac{(1/\delta)(1 - j\omega C_{JC}/g_m)}{1 + j\omega[(C_{JE} + C_{JC})/g_m\delta + \tau_F/\delta]} \approx \left(\frac{\tau_{BF}}{\tau_F}\right)\left[1 + j\omega\left(\tau_{BF} + \frac{(C_{JE} + C_{JC})\tau_{BF}}{g_m\tau_F}\right)\right]^{-1} \tag{7.5.14}$$

式中利用了等式(7.4.4),并忽略了分子中与频率有关的项,因为只有频率高到分母中的虚数项占支配地位时,这一项的作用才需要考虑。

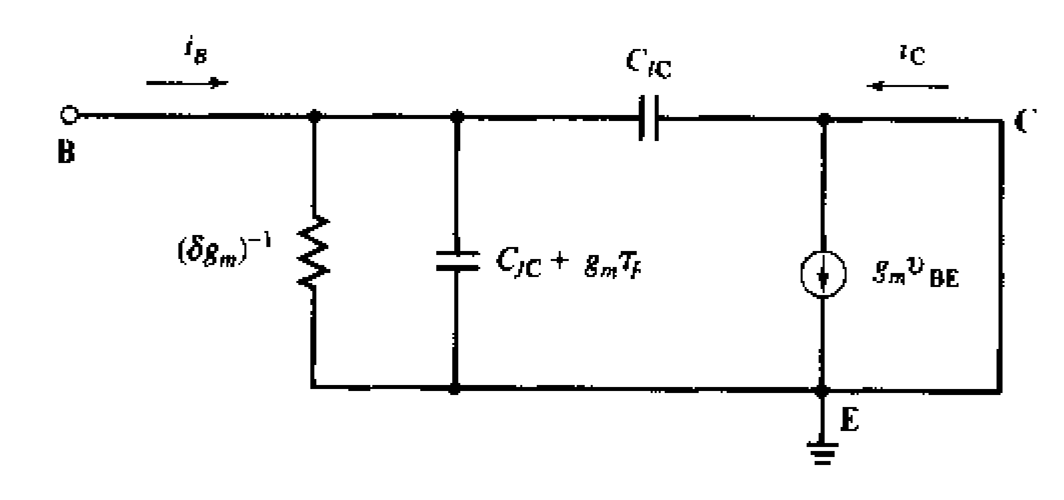


图 7.30 求 f_T 和 τ_F 关系的等效电路

随着频率的增加,电流增益下降,并且由等式(7.5.14)知,电流增益下降至 1 的频率 f_T 近似为

$$f_T = \frac{1}{2\pi\left(1 + \frac{C_{jE} + C_{jC}}{g_m \tau_F}\right)\tau_F} \qquad (7.5.15)$$

用这个等式求解 τ_F,得到

$$\tau_F = \frac{1}{2\pi f_T} - \frac{(C_{jE} + C_{jC})}{g_m} \qquad (7.5.16)$$

这样,通过测量短路电流增益(即集电极交流短路的电流增益)与频率的关系,就可以得到电荷控制参数 τ_F。在对数坐标系中绘出增益与频率的关系,并外推到增益为 1,即可得到参数 f_T,然后就可以利用等式(7.5.16)求出 τ_F。低频电流增益 β_F 可用于计算 $\tau_{BF} = \beta_F\tau_F$(等式(7.4.4))。在反向放大区重复上述测量,可得到参数 β_R 和 τ_{BR}。

7.6　双极晶体管的频率限制

从前一节我们知道,随着频率的增加,高频电流增益下降,这是因为移动电荷到晶体管的不同部位需要一定的时间。对于放大电路,小信号增益通常是最重要的晶体管参数。频率与增益之间的关系可以用两个参数来体现:输出短路时,晶体管电流增益下降至 1 的频率 f_T(7.5 节介绍的短路截止频率);考虑了基区电阻的参数 f_{max}(最高振荡频率)。

利用电荷控制模型,即考虑晶体管不同区域电荷的移动和电荷重新分布,可以得到这些参数的比较有用的表达式。把可动电荷按与其相关的电流来划分,并考虑相应的延迟时间,那么总延迟时间 $\tau_{EC} = 1/(2\pi f_T)$ 可写成与发射区、基区、集电区相关的几部分之和[12-15]

$$\frac{1}{2\pi f_T} = \tau_{EC} = \tau_C + \tau_B + \tau_E \qquad (7.6.1)$$

其他成分,包括与各区电阻(实际电阻和"动态"电阻)和结电容对应的 RC 网络时间常数也需要考虑进来,这些成分我们将简要分析。反映其他电荷移动的项将在更为精确的分析中考虑。

首先考虑第一项基区渡越时间 τ_B(见 7.3 节),它描述了基区少子从准中性基区的发射结边界运动到集电结边界所需要的时间,以及包围中性基区的耗尽区边界响应电压改变的充放电所需的时间(即发射结和集电结电容的充电时间)。

对于均匀基区 npn 晶体管,通过准中性基区的电流主要是扩散流,基区渡越时间可用等式(7.3.3)求出,我们重复在下面的公式中:

$$\tau_B = \frac{Q_B}{I_C} = \frac{qx_B n_{po}/2}{qD_{nB}n_{po}/x_B} = \frac{x_B^2}{2D_{nB}} \qquad (7.6.2)$$

式中 x_B 为准中性基区宽度,D_{nB} 为基区少子扩散系数。对于非均匀基区晶体管,基区电场将加速少子通过基区,漂移项将是非常重要的输运机制。当漂移项起主要作用时,基区渡越时间为

$$\tau_B = \int_0^{x_B}\frac{dQ}{I_C} = \int_0^{x_B}\frac{qn_p(x)dx}{qn_p(x)\mu_n\mathscr{E}} = \frac{x_B}{\mu_n\mathscr{E}} \qquad (7.6.3)$$

式中 μ_n 为基区少子迁移率,$\mathscr{E}$ 为基区电场(这里的推导假设 $\mathscr{E}$ 为常数)

可以通过杂质分布梯度或改变基区带隙来建立基区电场。如果基区掺杂浓度指数分布,

那么等式(4.1.14)和(4.1.15)描述的电场为常数

$$\mathscr{E}=\frac{kT}{qx_B}\ln\left(\frac{N_{BE}}{N_{BC}}\right) \tag{7.6.4}$$

式中 N_{BE} 和 N_{BC} 分别为准中性基区发射结边界和集电结边界处的掺杂浓度。

为了获得高电场,准中性基区发射结边界的掺杂浓度越高越好,集电结边界的掺杂浓度越低越好。发射结边界的基区掺杂浓度受到发射结注入效率的限制,典型值为 $5\times10^{17}\text{cm}^{-3}$ 范围。为了减轻 Early 效应,集电结边界的基区掺杂浓度必须高于集电区掺杂浓度,使集电结空间电荷区主要向集电区扩展。集电区掺杂浓度的典型值为 $N_{dC}=10^{16}\text{cm}^{-3}$,集电结边界的基区掺杂浓度至少为 10^{17}cm^{-3}。因此,通过掺杂分布获得的最大电场强度受限于相邻的发射区和集电区的掺杂浓度。如果基区宽度为 100nm,掺杂浓度如前面所述,那么最大电场强度 $\mathscr{E}\leq 1\times10^4\text{V/cm}$。可以通过改变基区的带隙来获得更大的电场强度,如图 4.2(b)所示。对于宽度为 100nm 的基区,线性改变带隙,改变量 $\Delta E_g=0.2\text{eV}$,那么电场强度 $\mathscr{E}=\Delta E_g/(qx_B)=2\times10^4\text{V/cm}$。减小基区宽度,电场的增加将更为显著。

对于很薄的基区,基区渡越时间非常小,那么限制器件速度的主要因素是其他的延迟时间。首先,载流子通过宽度为 x_{dC} 的集电结耗尽区的时间有限。对晶体管合理设计和偏置,那么载流子通过集电结的极限速度为 v_l(也称为饱和速度 v_{sat}),集电结渡越时间为

$$\tau_C=\frac{x_{dC}}{2v_{sat}} \tag{7.6.5}$$

(分母中的因子 2 是由于假设约有一半的电荷超前通过基区[16,17])。可以通过增大集电区掺杂浓度来减小 τ_C。然而,集电区掺杂浓度的增加会降低集电结击穿电压,因此需折衷考虑。

由 7.4 节讨论的电荷控制模型知,可以将基区渡越时间看做是改变注入到基区的电荷的分布而导致的延迟,可由电荷 Q_B 除以电子电流 I_{nB} 给出。类似的延迟时间是与从基区注入到发射区的空穴的再分布相关的。这个不需要的基极电流成分(见 6.2 节)降低了晶体管的频率响应,同时也降低了电流增益。与准中性发射区电荷存储效应有关的延迟时间可以写成发射区空穴电荷除以充电电流。考虑到具有合理的发射效率的晶体管,充电电流主要是电子电流,并且 $J_{nB}\approx\beta J_{pE}$,那么

$$\tau_E=\frac{Q_{pE}}{\beta J_{pE}}=\frac{x_E^2}{2D_{pE}}\frac{GN_B}{GN_E} \tag{7.6.6}$$

式中 x_E 为准中性发射区厚度,GN 为等式(6.2.3)定义的 Gummel 数。我们还可以假设 $\beta\approx GN_E/GN_B$,少子扩散长度 L_{pE} 远大于 x_E。(更完善的分析表明,重掺杂发射区的带隙变窄将导致 τ_E 增加。)

除 τ_C、τ_B 和 τ_E 以外,还需要考虑与 RC 乘积有关的延迟。我们可以将这些乘积分成两项,一项与势垒电容有关

$$\tau_1=r_eC_{jE} \tag{7.6.7}$$

式中 r_e 为发射区"动态"电阻 $(dI_C/dV_{BE})^{-1}=kT/qI_C=1/g_m$。另一项为

$$\tau_2=(r_e+R_E+R_C)C_{jC} \tag{7.6.8}$$

式中 R_E 和 R_C 分别为准中性发射区和集电区的与结构有关的电阻与电极接触电阻之和。那么

$$\frac{1}{2\pi f_T} = \tau_{EC} = \tau_r + \tau_B + \tau_E + \tau_1 + \tau_2 \tag{7.6.9}$$

电流较低时，τ_r 较大，是限制 f_T 的主要因素。随着电流的增大，τ_r 减小，其他项开始占支配地位。高电流时，Kirk 效应展宽了基区，致使 τ_B 增大，f_T 减小；其他的大电流效应也会降低 f_T。利用 $1/f_T$ 与 $1/I_E$ 关系曲线，可以找出 τ_r 起主要作用的区域以及其他效应起主要作用的区域，以了解限制频率响应的物理机构。

尽管 f_T 被广泛用于描述双极晶体管的频率响应，但它没有考虑基区电阻 R_B 的限制。另一个参数，即最高振荡频率 f_{max}（功率增益为1的频率），考虑了基区电阻 R_B 的影响。因此，f_{max} 更为全面地反映了晶体管的频率特性。

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_B C_{jC}}} \tag{7.6.10}$$

等式(5.4.2)、等式(5.4.4)以及少子和多子关系式 $p_{nE} = n_i^2/N_{dE}$ 表明，从基区注入到发射区的空穴取决于基区掺杂浓度。为降低基区电阻，必须提高基区掺杂浓度。然而，这将增加注入到发射区的空穴，增大了延迟时间 τ_E，同时也降低了发射效率和晶体管电流增益。

异质结晶体管由于发射结能带的不连续，减弱了基区电阻和频率响应之间的折衷考虑。对于相同的集电极电流（通过发射结注入的电子电流），由于降低了电子势垒，异质结晶体管需要的 V_{BE} 更低。低工作电压降低了注入到发射区的空穴流，也降低了发射区的电荷存储效应。能带的不连续产生的阻止空穴从基区注入到准中性发射区的因子为 $\exp(\Delta E_g/kT)$，因此发射区延迟时间 τ_E 降低至（利用等式(7.6.6)）

$$\tau_E = \frac{x_E^2}{2D_{pE}}\frac{GN_B}{GN_E}\exp\left(\frac{-\Delta E_g}{kT}\right) \tag{7.6.11}$$

这样，与同质结相比，能带不连续的异质发射结减弱了 τ_E 与基区电阻之间的折衷。异质结允许我们在固定基区掺杂浓度的情况下降低延迟时间。另一种选择是在保持 f_T 不变的情况下，大幅度提高基区掺杂浓度以提高 f_{max}。通过适当地改变参数，可以得到 f_T 和 f_{max} 的最优值。（注意 HBT 和同质结晶体管的 τ_C 改变不大，这两种晶体管的集电结耗尽区主要都扩展在中等掺杂的集电区，其掺杂浓度的限制条件都是击穿电压。）

例题　双极晶体管的频率响应

利用合理的材料参数，计算硅同质结双极晶体管以及发射结处价带能量差为0.1eV的异质结双极晶体管的 f_T 和 f_{max}（利用硅材料性质求解其他参数）。要求晶体管的增益 $\beta = 200$，发射极-集电极击穿电压 $BV_{CEO} = 7V$。晶体管的直流偏置为 $I_C = 10mA$，$V_{CB} = 3V$。

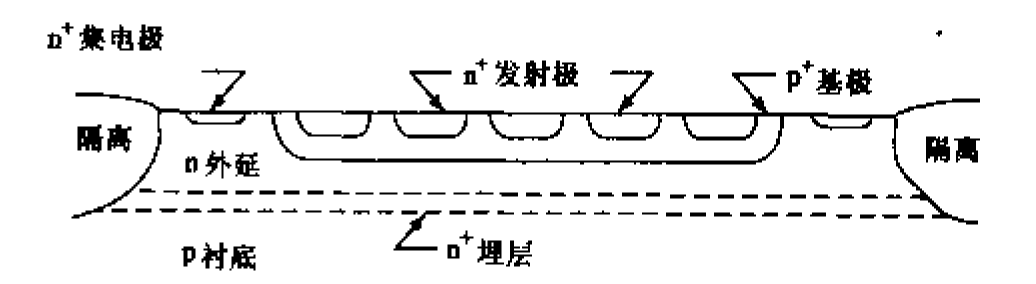
解：利用题中的设计标准，可求出集电结击穿电压 $BV_{CBO} = BV_{CEO}\beta^{1/m}$［式中 $m \sim 4$（等式(6.5.1)）］必须为26V，相应的集电区掺杂浓度 N_{dC} 为 $3\times10^{16}cm^{-3}$。已知 $V_{CB} = 3V$，那么集电结耗尽区宽度为0.4μm，利用等式(7.6.5)求出 $\tau_C = 2.0ps$。

为降低同质结晶体管的基区渡越时间，基区掺杂必须为递降分布。如果基区宽度为

100nm,电子迁移率为 $400\text{cm}^2/\text{V}-\text{s}$,电场为 10kV/cm,那么基区渡越时间 $\tau_B=2.5\text{ps}$。(相同宽度的基区均匀掺杂晶体管,基区渡越时间大约为其两倍。)

如果中性发射区宽度 $x_E=0.25\mu\text{m}$,空穴扩散系数为 $3\text{cm}^2/\text{s}$、比值 GN_E/GN_B 为 200(假设电流增益由发射区和集电区 Gummel 数的比值决定),利用等式(7.6.6)可计算出发射区延迟 $\tau_E=0.5\text{ps}$。对于 10mA 的集电极电流,发射极动态电阻 $r_e=2.6\Omega$。

f_T表达式中的其他参数取决于晶体管的几何结构。如下图所示,假设多发射极晶体管有 5 个发射极条,每条长 $20\mu\text{m}$,宽 $1\mu\text{m}$,总发射区面积为 $100\mu\text{m}^2$。利用这些尺寸,发射极电阻 R_E(发射区体电阻加上接触电阻)的合理值为 1Ω。



发射结耗尽区电容 C_{jE}为 200fF 数量级($1\text{fF}=10^{-15}\text{F}$)[考虑了侧壁电容后,$C_{jE}$实际上更高一些(可能为其两倍)]。集电结耗尽区电容 C_{jC}的估算可以利用前面求出的集电结耗尽区宽度为 $0.4\mu\text{m}$,并假设集电区面积为发射区面积的 3 倍,那么 $C_{jC}=80\text{fF}$。

集电区串联电阻取决于版图的详细资料以及制造电极和埋层的工艺,其组成为:外延层未耗尽部分的电阻,埋层的横向电阻(考虑到每个发射极电极接触的位置进行合理地加权),埋层和集电极之间的纵向电阻,集电极接触电阻。根据我们所用的数据,可以估计集电极电阻为 7Ω(埋层电阻为 2Ω,电极接触加上埋层和电极之间的纵向电阻为 5Ω)。

利用这些数据,由等式(7.6.1)、(7.6.7)和(7.6.8)可估算出 f_T(所有项的单位均为 ps)

$$\begin{aligned}\tau_{EC} &= \tau_C + \tau_B + \tau_E + (r_e + R_E + R_C)C_{jC} + r_e C_{jE}\\ &= 2.0 + 2.5 + 0.5 + (2.6 + 1 + 7)\times 80\times 10^{-15} + 2.6\times 210\times 10^{-15}\\ &= 2.0 + 2.5 + 0.5 + 0.9 + 0.5\\ \tau_{EC} &= 6.4\ \text{ps}\end{aligned}$$

相应地,$f_T=(2\pi\tau_{EC})^{-1}=2.5\times10^{10}\text{Hz}=25\text{GHz}$。对于这个设计,主要的延迟发生在集电结耗尽区和基区。前者受到的限制是要求足够的击穿电压,后者受到的限制是需要降低基区电阻以获得较高的 f_{max}。

计算 f_{max}需要知道基区电阻,基区电阻的组成为:发射区下方的本征基区电阻,有源区和电极之间的非本征基区电阻,电极接触电阻。根据我们所用的几何尺寸,这些参数的合理数值为:本征基区电阻为 8Ω,非本征基区电阻为 2Ω,电极接触电阻为 2Ω,总基区电阻为 12Ω。将基区电阻和上述其他参数的数值代入等式(7.6.10),得到 $f_{max}=32\text{GHz}$。

对于异质结双极晶体管,我们主要考虑 f_T和 f_{max}的提高。就本例题而言,为提高 f_T和 f_{max},我们可以将基区宽度减小为原来的 1/2 以改善基区渡越时间和 f_T,同时将基区掺杂浓度提高为原来的 10 倍以降低基区电阻,提高 f_{max}。发射结的异质结势垒允许我们在增大基区掺杂浓度的同时,不会导致向发射区注入空穴的增加以及发射区延迟时间 τ_E增大。本例

题中，基区 Gummel 数的增大因子为5，但势垒降低 0.1eV 可导致向发射区的注入下降 46 倍。因此 τ_E 大约降至原来的 1/10，约为 0.05ps，该数值可忽略。

经过上述的改进，$\tau_{EC}=2.0+1.25+0.05+0.9+0.5=4.7\text{ps}$，$f_T=34\text{GHz}$，与同质结晶体管的 25 GHz 相比有显著的提高。f_{max} 的提高更为明显，基区电阻的本征部分下降为原来的 1/5，如果假设非本征基区电阻和电极接触电阻不变，那么总基区电阻从 12Ω 降至 5.6 Ω。利用等式(7.6.10)，代入新的 f_T 和 R_B 值，得到 $f_{max}=55\text{GHz}$，而同质结晶体管的 f_{max} 为 32 GHz。本例题中选择不同的参数可以使我们主要提高 f_T 的值，而不是 f_{max} 的值。

7.7 计算机模拟中的双极晶体管模型†

计算机模拟优先考虑的是精度，而不是概念和计算简便。晶体管的计算机模型要求能准确描述晶体管的大信号和小信号工作，器件特性易于用模型参数表征，同时要求模型参数易于获得和验证。以6.4节介绍的 EM 方程为基础的模型能较好地满足上述要求[18]。

我们讨论的切入点为"传输型"EM 方程，由等式(6.4.2)和等式(6.4.3)组成。该方程组根据发射结和集电结之间的连接电流以及发射结和集电结二极管电流分量来描述晶体管。

等式(6.1.14)给出了连接电流密度 J_n 的表达式，那么总连接电流为

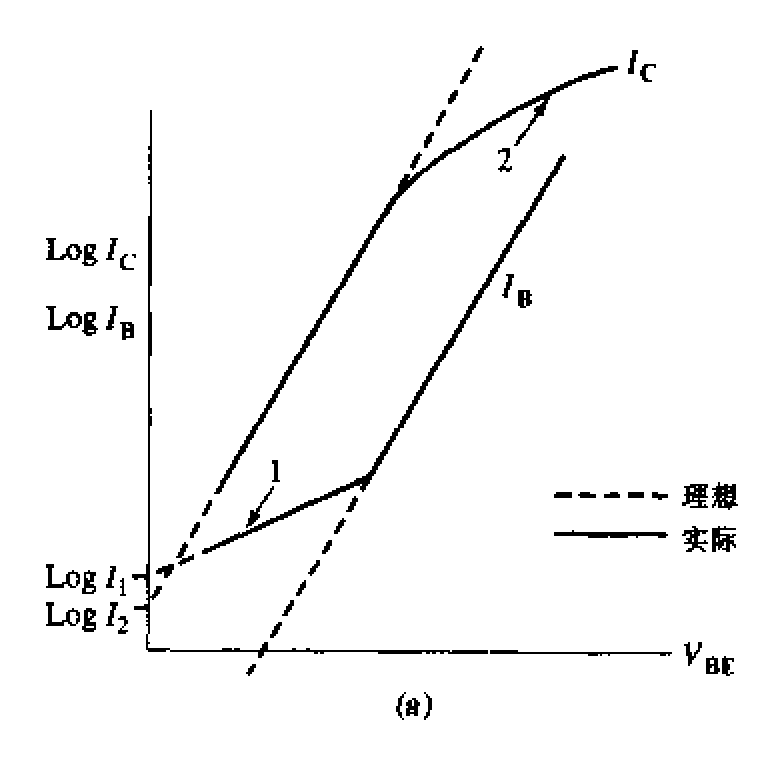
$$I_n=I_S\left[\exp\left(\frac{V_{BC}}{V_t}\right)-\exp\left(\frac{V_{BE}}{V_t}\right)\right] \tag{7.7.1}$$

利用该等式以及等式(6.4.7(a))和(6.4.7(b))，EM 方程可以写成如下形式

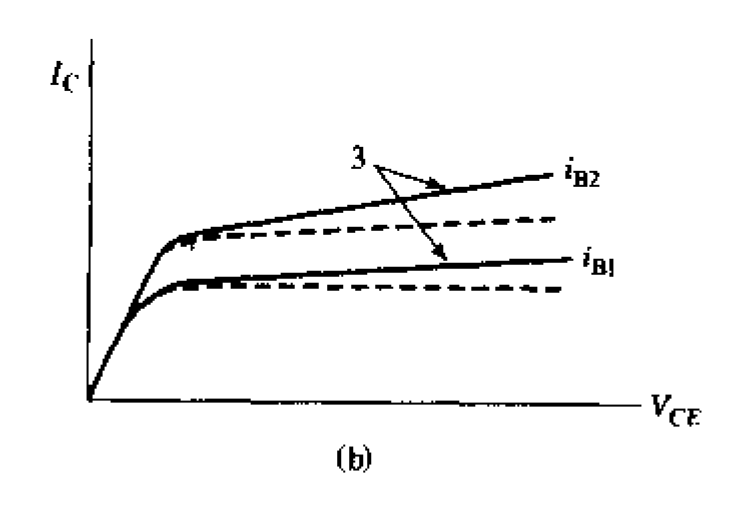
$$\begin{aligned}I_C&=-I_n-\frac{I_S}{\beta_R}\left[\exp\left(\frac{V_{BC}}{V_t}\right)-1\right]\\I_E&=I_n-\frac{I_S}{\beta_F}\left[\exp\left(\frac{V_{BE}}{V_t}\right)-1\right]\\I_B&=\frac{I_S}{\beta_F}\left[\exp\left(\frac{V_{BE}}{V_t}\right)-1\right]+\frac{I_S}{\beta_R}\left[\exp\left(\frac{V_{BC}}{V_t}\right)-1\right]\end{aligned} \tag{7.7.2}$$

该方程组用三个参数 I_S、β_F 和 β_R 来描述基本的 EM 关系。EM 模型中没有反映的效应，例如本章前面所描述的效应，可以通过在方程组(7.7.2)中加入其他项来体现。Gummel 和 Poon[19] 提出了一种相对更直接的方法，通过修正方程组(7.7.2)，就可以描述晶体管的三个重要的二阶效应：(1)低发射结偏置时的发射结空间电荷区复合；(2)大电流时电流增益的下降；(3)发射结和集电结之间的连接电流引起的基区宽度调变效应(Early 效应)。这些二阶效应导致了图7.31(a)和7.31(b)所示的电流偏离理想特性的现象。考虑了这些二阶效应的 Gummel-Poon 模型(GP 模型)更适合于晶体管的计算机模拟。

空间电荷区的复合 由第5章知，空间电荷区的复合将影响二极管的电流与电压的关系。EM 模型模拟该效应的方法是增加四个参数，将基极电流定义为理想成分和非理想成分的叠加新增加的参数 I_1、I_2、n_e 和 n_c 在低基极-发射极偏置时测量。例如，I_1 可通过截取一段 $\log I_B-V_{BE}$ 曲线，外推至 $V_{BE}=0$ 获得(见图7.31)。



(a)



(b)

图7.31 工作在正向放大区的双极晶体管的二阶效应。图中的数字与正文中的数字一一对应。由等式(7.7.3)知,外推到基极-发射极电压为零时的基极电流为 I_1

$$I_B = \frac{I_S}{\beta_F}\left[\exp\left(\frac{V_{BE}}{V_t}\right) - 1\right] + I_1\left[\exp\left(\frac{V_{BE}}{n_e V_t}\right) - 1\right] + \frac{I_S}{\beta_R}\left[\exp\left(\frac{V_{BC}}{V_t}\right) - 1\right] + I_2\left[\exp\left(\frac{V_{BC}}{n_c V_t}\right) - 1\right] \tag{7.7.3}$$

Early 效应和大电流效应 大电流效应(2)和 Early 效应(3)可以通过修正连接电流的系数 I_S体现。从6.1节知,J_S与基区多子电荷浓度 Q_B成反比。如果改写等式(6.1.13)和等式(6.1.15)来表示基区多子电荷总量 Q_{BT}和总饱和电流 I_S,得到

$$I_S = J_S A_E = \frac{q^2 A_E^2 n_i^2 \tilde{D}_n}{Q_{BT}} \tag{7.7.4}$$

式中

$$Q_{BT} = qA_E \int_0^{x_B} p(x)\,dx \tag{7.7.5}$$

GP 模型中,Q_{BT}是与偏置有关并易于计算的量。首先,基区"内建"多子电荷(平衡态基区多子电荷)Q_{B0}为

$$Q_{BO} = qA_E \int_0^{x_B} N_a(x)\,dx \tag{7.7.6}$$

此外，还需要考虑发射结和集电结空间电荷区存储的电荷（Q_{JE} 和 Q_{JC}），以及正向和反向注入到基区的少子电荷，那么基区多子总电荷 Q_{BT} 为

$$\begin{aligned}Q_{BT} = Q_{BO} + C_{JE}V_{BE} + C_{JC}V_{BC}\frac{A_E}{A_C} + \frac{Q_{BO}}{Q_{BT}}\tau_F I_S\left[\exp\left(\frac{V_{BE}}{V_t}\right)-1\right] \\ + \frac{Q_{BO}}{Q_{BT}}\tau_R I_S\left[\exp\left(\frac{V_{BC}}{V_t}\right)-1\right]\end{aligned} \tag{7.7.7}$$

通过定义几个参数，可以将等式(7.7.7)写成更易于处理的形式

$$\begin{aligned} q_b &\equiv \frac{Q_{BT}}{Q_{BO}}; \qquad I_{KF} \equiv \frac{Q_{BO}}{\tau_F}; \\ I_{KR} &\equiv \frac{Q_{BO}}{\tau_R}; \qquad |V_A| \equiv \frac{Q_{BO}}{C_{JC}}\frac{A_C}{A_E} \\ |V_B| &\equiv \frac{Q_{BO}}{C_{JE}} \end{aligned} \tag{7.7.8}$$

等式(7.7.8)将基区总电荷 Q_{BT} 用 Q_{BO} 归一化，这个无量纲的变量称为归一化的基区多子电荷 q_b。两个电荷控制时间常数 τ_F 和 τ_R 与 Q_{BO} 一起定义了"膝点电流" I_{KF} 和 I_{KR}。Early 电压 V_A 以及反向工作 Early 电压 V_B 的定义与等式(7.1.8)意义相同。

利用归一化参数，等式(7.7.7)可写成

$$q_b = q_1 + \frac{q_2}{q_b} \tag{7.7.9}$$

式中 q_1 和 q_2 为辅助变量，其定义为

$$\begin{aligned} q_1 &= 1 + \frac{V_{BE}}{|V_B|} + \frac{V_{BC}}{|V_A|} \\ q_2 &= \frac{I_S}{I_{KF}}\left[\exp\left(\frac{V_{BE}}{V_t}\right)-1\right] + \frac{I_S}{I_{KR}}\left[\exp\left(\frac{V_{BC}}{V_t}\right)-1\right] \end{aligned} \tag{7.7.10}$$

这些新增加的变量是为了更方便地指出二阶效应的作用。如果 Early 效应可以忽略，那么 q_1 将趋于 1。如果大电流效应不重要，那么 q_2 将非常小。

这样，基区宽度调制效应可以通过两个 Early 电压来建模，而大电流效应通过膝点电流 I_{KF} 和 I_{KR} 来描述。GP 模型需要定义的参数包括基本 EM 模型的三个参数 I_S、β_F 和 β_R，以及描述空间电荷区复合效应的四个参数 I_1、I_2、n_e 和 n_c。EM 模型参数在中等电流区有效，在大电流区没有意义。

最后，用归一化基区多子电荷 q_b 模拟基区宽度调变和多子电荷调制效应，q_b 由参数 I_{KF}、I_{KR}、V_A 和 V_B 确定。完整的模型需要 11 个参数以及温度参数（用于计算 V_t）。npn 晶体管完整的模型方程为

$$I_B = \frac{I_S}{\beta_F}\left[\exp\left(\frac{V_{BE}}{V_t}\right) - 1\right] + I_1\left[\exp\left(\frac{V_{BE}}{n_e V_t}\right) - 1\right] + \frac{I_S}{\beta_R}\left[\exp\left(\frac{V_{BC}}{V_t}\right) - 1\right] + I_2\left[\exp\left(\frac{V_{BC}}{n_c V_t}\right) - 1\right]$$

$$I_C = \frac{I_S[\exp(V_{BE}/V_t) - \exp(V_{BE}/V_t)]}{q_b} - \frac{I_S}{\beta_R}\left[\exp\left(\frac{V_{BC}}{V_t}\right) - 1\right] - I_2\left[\exp\left(\frac{V_{BC}}{n_c V_t}\right) - 1\right] \qquad (7.7.11)$$

$$q_b = \frac{q_1}{2} + \frac{\sqrt{q_1^2 + 4q_2}}{2}$$

$$q_1 = 1 + \frac{V_{BE}}{|V_B|} + \frac{V_{BC}}{|V_A|}$$

$$q_2 = \frac{I_S}{I_{KF}}\left[\exp\left(\frac{V_{BE}}{V_t}\right) - 1\right] + \frac{I_S}{I_{KR}}\left[\exp\left(\frac{V_{BC}}{V_t}\right) - 1\right]$$

为说明这组方程的有效性,我们考虑工作在正向放大区的小注入情形下的晶体管,可求出 $q_2 \simeq 0$(因为图 7.32 中的集电极电流远小于膝点电流 I_{KF})。在这种情形下,有

$$I_C \simeq \frac{I_S \exp(V_{BE}/V_t)}{1 + V_{BC}/|V_A|} \simeq I_S \exp\left(\frac{V_{BE}}{V_t}\right)\left(1 - \frac{V_{BC}}{|V_A|}\right) \qquad (7.7.12)$$

和

$$\frac{\partial I_C}{\partial V_{CB}} = \frac{I_C}{|V_A|} \qquad (7.7.13)$$

上述等式与等式(7.1.3)一致。

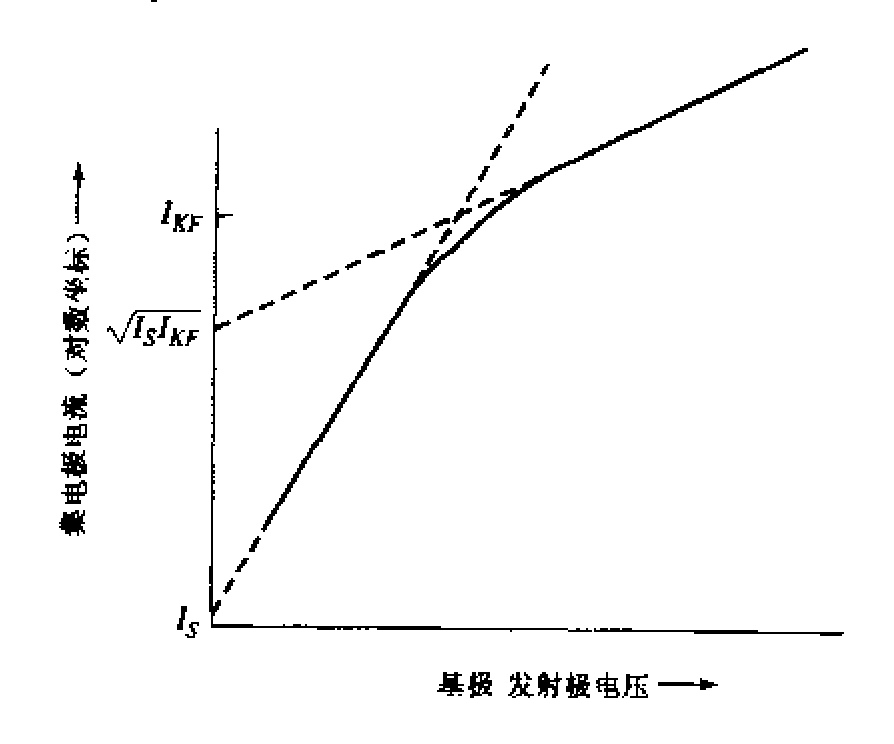


图 7.32 半对数坐标系中集电极电流与 V_{BE} 的关系,用于说明大电流情形下的 GP 模型。小电流区和大电流区的渐近线相交于"膝点"电流 $I_C = I_{KF}$

大电流条件下,Early 效应远小于大注入效应,$q_2 > q_1$,归一化的基区电荷 q_b 的渐近线与大电流区渐近线一致

$$q_b = \sqrt{\frac{I_S}{I_{KF}}}\exp\left(\frac{V_{BE}}{2V_t}\right) \tag{7.7.14}$$

因此，集电极电流为（见图 7.32）

$$I_C = \sqrt{I_S I_{KF}}\exp\left(\frac{V_{BE}}{2V_t}\right) \tag{7.7.15}$$

以下问题作为习题留下：证明“膝点”电流 $I_C = I_{KF}$ 处是小电流区 GP 模型渐近线与大电流区渐近线（等式（7.7.15））的交点。等式（7.7.15）成立的物理基础是基区已出现足够高的大注入效应，已经导致基区多子浓度与偏置有关，正如等式（7.2.3）所描述的。

7.8 器件：pnp 双极晶体管

如果集成电路工艺中只有一种类型的晶体管（例如 npn 或 pnp），那么制造工艺将会非常简单，并能降低成本，提高成品率。由于电子的迁移率比空穴高，npn 晶体管的使用更为广泛，尤其是用于如图 7.33 所示的驱动晶体管。完整的电路还需要一个负载元件接在 npn 晶体管的集电极和电源 V_{CC} 之间。尽管可以用电阻作为负载，但用晶体管更合适。晶体管负载的传输特性可以使电路获得更快的开关速度。另外，与电阻相比，IC 晶体管可提高集成度（降低成本）。

7.8.1 衬底 pnp 晶体管

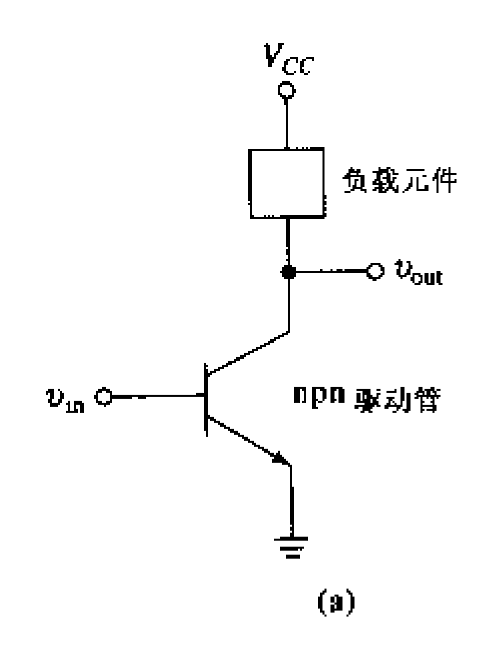
npn 晶体管性能较好的原因之一是晶体管作用发生的区域远离表面，并且结的大部分区域均匀且与表面平行。用标准的平面工艺获得具有这些特点的 pnp 晶体管的方法是将 p 型扩散区（npn 晶体管的基区）作为 pnp 晶体管的发射区，图 7.34 为其横截面示意图。

衬底 pnp 晶体管用外延层做基区，外延层和衬底之间的 pn 结作为集电结。由于集电区是集成电路中的衬底，pnp 晶体管的集电区不能相互隔离。因此，这种衬底 pnp 晶体管只能用在集电极交流接地的电路，比如射极跟随器电路。衬底 pnp 晶体管对许多集成电路是有用的，但显然不能满足要用到 pnp 晶体管的各种电路的需求。尽管衬底 pnp 晶体管没有类似双扩散 npn 晶体管的基区内建电场，它的 β 仍能达到 100（电流为 1mA 时）。根据具体的工艺，纵向 pnp 晶体管工作范围为 1μA 到 10mA，f_T 值高达 50 MHz。

7.8.2 横向 pnp 晶体管

用标准工艺也可制造出集电区相互隔离的 pnp 晶体管，如图 7.35 所示，用两个靠得很近的 p 型扩散区（标准 npn 晶体管的基区）作为发射区和集电区。这种晶体管称为横向 pnp 晶体管，因为晶体管作用区域是横向的，也就是说，发射区和集电区之间的区域平行于表面。这种设计牺牲了晶体管作用远离表面的优点。因此，横向 pnp 晶体管的性能明显不如标准的纵向 npn 晶体管。尽管如此，横向 pnp 晶体管仍广泛用于模拟和数字集成电路中。

图 7.35 所示的横向 pnp 晶体管的两个集电极被连接在一起。为了提高电流增益，典型的做法是让集电区完全环绕发射区。横向 pnp 晶体管同样需要埋层，埋层可从两个方面提高电



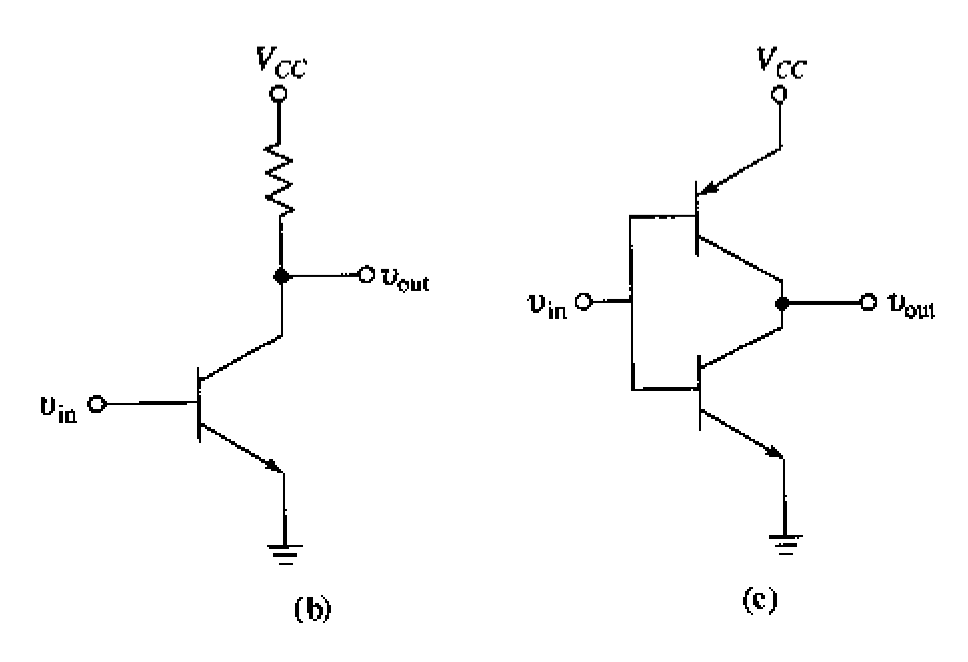


图 7.33 (a)双极晶体管电路中的负载元件;
(b)电阻用做负载;(c)互补 pnp 晶体管用做负载

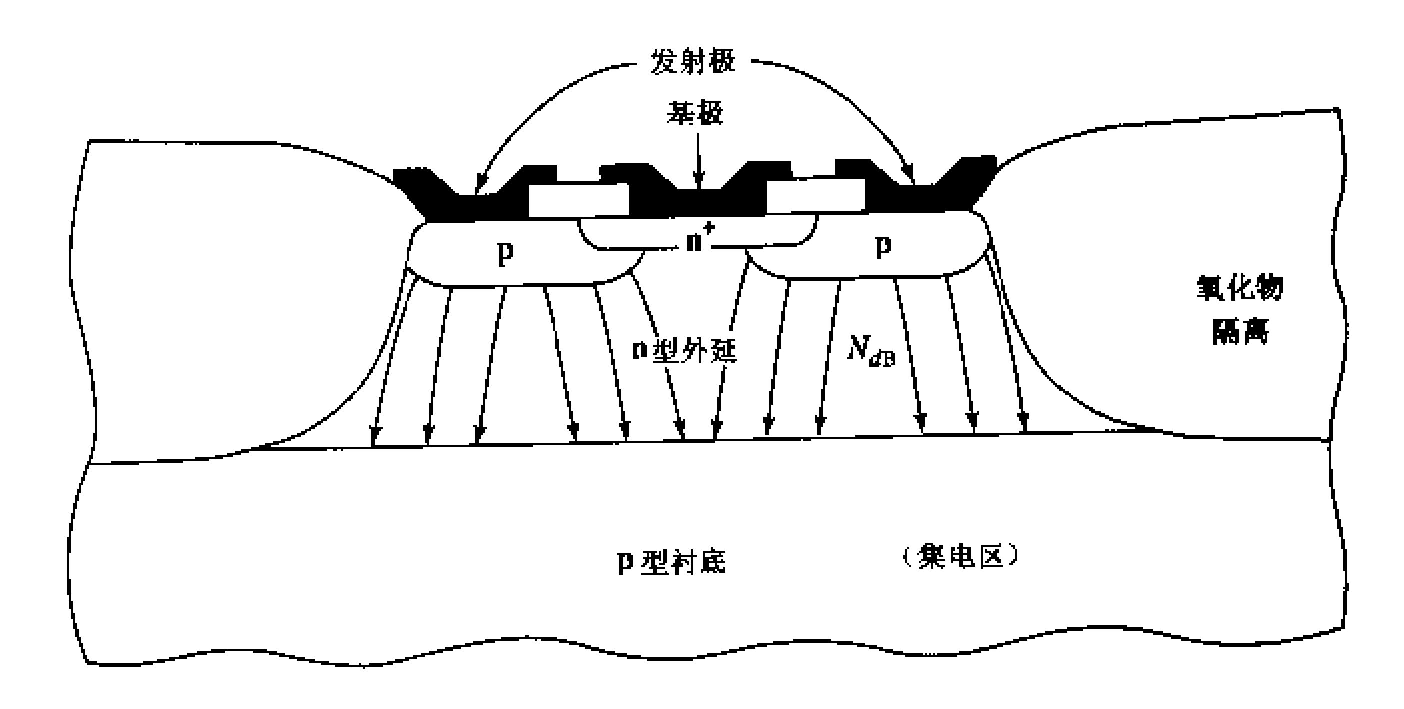


图 7.34 衬底 pnp 晶体管横截面示意图。图中的流线为连接电流

流增益和频率响应特性:(1)降低了基区电阻;(2)抑制外延层和衬底形成的结对空穴的收集。第二点可以这样理解,埋层的杂质分布梯度形成的内建电场阻挡了空穴从外延层进入衬底。换句话说,我们可以认为埋层提高了寄生衬底 pnp 晶体管的基区掺杂 Q_B,由等式(6.1.15)知,

这将减小从寄生器件流失的空穴,因此提高了横向 pnp 晶体管的电流增益。

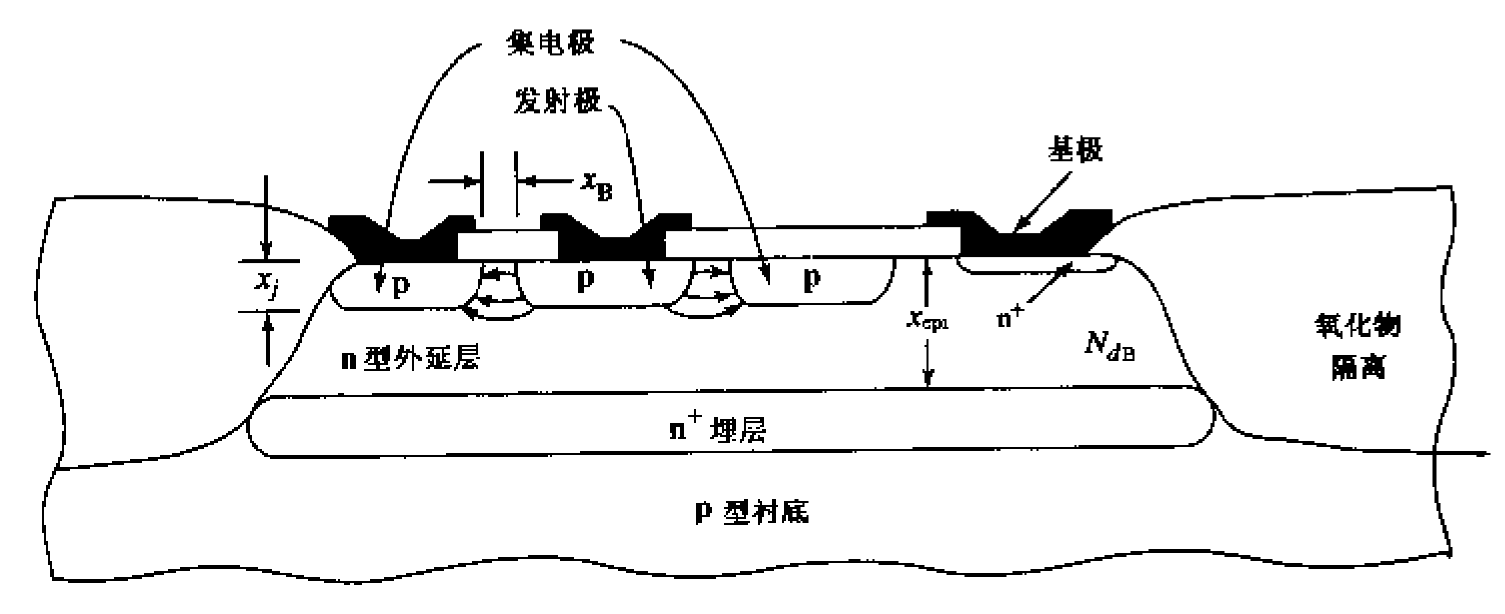


图 7.35　典型的横向 pnp 晶体管横截面示意图。扩散形成的集电区完全环绕发射区。图中电流线为连接电流

集电极电流　如图 7.35 所示,横向 pnp 晶体管的连接电流在发射结和集电结之间二维流动。由于外延基区均匀掺杂,在给定的发射极-基极偏置下,基区中沿发射结边缘注入的空穴浓度值都相同。载流子浓度梯度导致基区中的空穴向集电结扩散,其中靠近表面处浓度梯度最大,这是因为越靠近表面,基区宽度越小。随着离开表面距离的增加,两个扩散 p 型区之间的距离逐渐增加,浓度梯度也逐渐下降。因此,横向 pnp 晶体管的基区宽度只能用表面处的基区宽度近似(图 7.35 中的 x_B),并且通过发射结的空穴电流也不是均匀的。另外,连接电流还会受到外延层厚度和埋层结构的影响。

文献[20]详细分析了横向 pnp 晶体管的这些效应,并给出了等效求解连接电流 I_p 的经验方法:将晶体管等效成基区宽度为 x_B,发射结面积为 $P_E x_j$ 的一维晶体管,其中 P_E 为发射区的周长,x_j 为扩散结结深。那么,连接电流为

$$I_p = F\frac{qP_E x_j D_p n_i^2}{N_{dB} x_B}\exp\left(\frac{qV_{EB}}{kT}\right) \tag{7.8.1}$$

式中 F 取决于两个无量纲的比值:x_{epi}/x_j 和 x_B/x_j[20]。如图 7.35 所示,x_j 为发射区和集电区的扩散结深,x_{epi} 为埋层上方的外延层厚度,x_B 为硅-二氧化硅界面处发射区和集电区的距离。图 7.36 是摘自文献[20]的 $F(x_{epi}/x_j, x_B/x_j)$ 曲线。一种典型情况是两个变量的值都为 2,由图 7.36 知,F 约为 1.8。因此,总电流大约为一维分析预期值的两倍。

距离 x_B 由光刻掩膜版上发射区到集电区的距离减去横向扩散距离决定。最小光刻间距的典型值为 0.3μm,横向扩散略小于纵向扩散,x_B 一般为 0.2μm 量级。尽管扩散越深,x_B 越小,可以提高增益和频率响应特性,但是会降低可重复性和可靠性。

外延层一般为轻掺杂($\sim 10^{16}\,\text{cm}^{-3}$),因此不太高的发射结偏置就能使基区进入大注入状态(见习题 7.9)。Chou[20] 已经指出,大注入通过三种方式降低横向 pnp 晶体管的电流增益:(1)减轻了发射结处非平衡空穴密度 $p'(0)$ 对基极-发射极电压的依赖关系,可以用与等式(7.2.3)类似的空穴表达式描述;(2)由于 Webster 效应(见 7.3 节),扩散系数 D_p 将显著变化;(3)由于基区和发射区电阻,实际加在发射结上的电压将降低。

基极电流　分析 npn 晶体管时,只需要考虑基极电流的三个成分就能得到准确的模型。

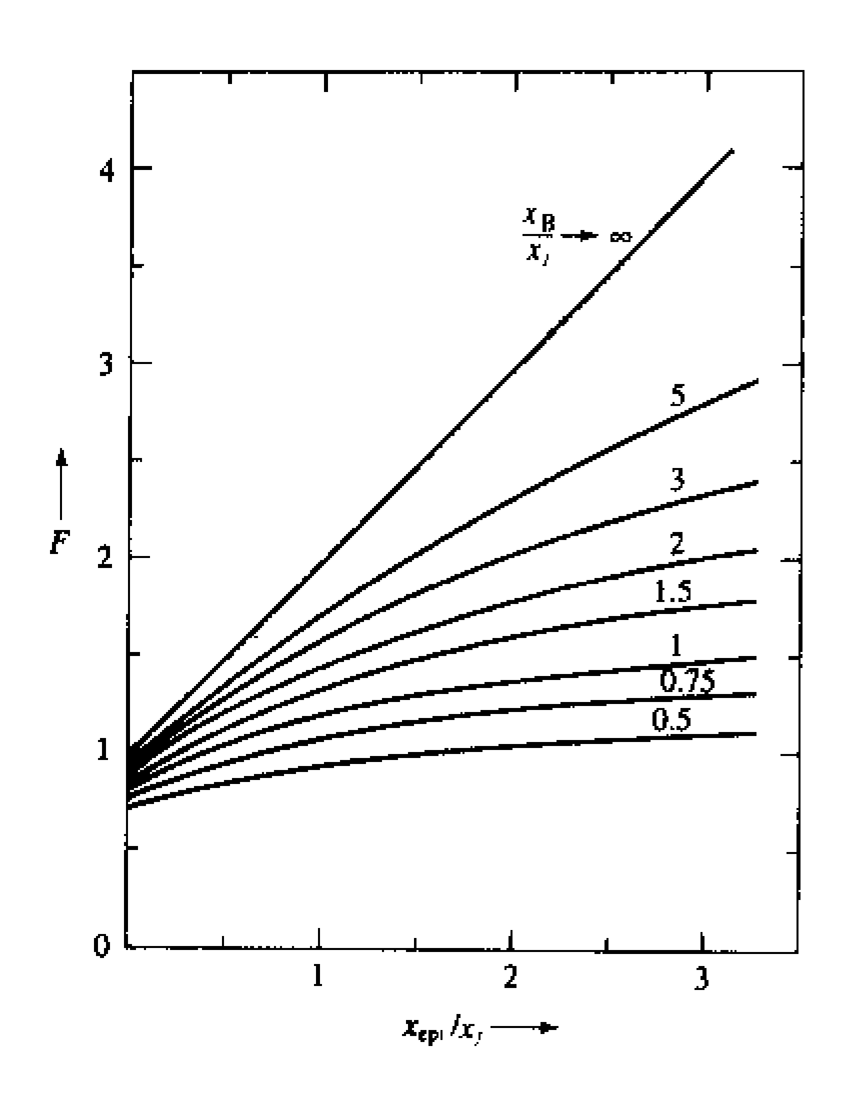


图 7.36　几何尺寸因子 F 表示实际晶体管集电极电流与一维模型获得的集电极电流之间的关系[20]

这三个成分是:向发射区注入的空穴电流(在大部分偏置条件下,该电流是 npn 晶体管基极电流的主要成分),基区复合电流,发射结空间电荷区复合电流(发射极电流较小时需要考虑)。横向 pnp 晶体管的基极电流主要也是这三个部分,只是电子和空穴的作用相互交换。然而,由于器件的硅-二氧化硅界面以及横向结构的影响,横向 pnp 晶体管还需要考虑其他两个重要的基极电流成分:硅-二氧化硅界面复合和埋层附近的复合。因此,横向 pnp 晶体管的基极电流总共包括五个部分,如图 7.37 所示。

与 npn 晶体管类似的三个基极电流成分可以用类似的等式描述,其中不同于表达式(6.2.4)的是基区复合电流。横向 pnp 晶体管基极电流的二维流动使这一项更为复杂。Chou[20] 给出的基区复合电流的表达式为

$$I_{B2}=\frac{qn_i^2[\exp(qV_{EB}/kT)-1]}{N_{dB}\tau_p}V_B^*\qquad(7.8.2)$$

式中 τ_p 为外延层空穴寿命,V_B^* 为由二等分基区宽度 x_B 的表面所定义的体积(图 7.37 中的阴影部分)。

如果定义埋层与未掺杂的外延层之间界面的复合速度为 s_{nn^+},则可以很方便地计算出流入埋层的空穴电流(I_{B4})。由等式(5.2.23)知,复合速度乘以注入的过剩载流子浓度为表面总复合率。假设外延层寿命相对较长,我们可以利用发射结过剩空穴的边界值写出

$$I_{B4}=qs_{nn^+}A_B^*\frac{n_i^2}{N_{dB}}\left[\exp\left(\frac{qV_{EB}}{kT}\right)-1\right]\qquad(7.8.3)$$

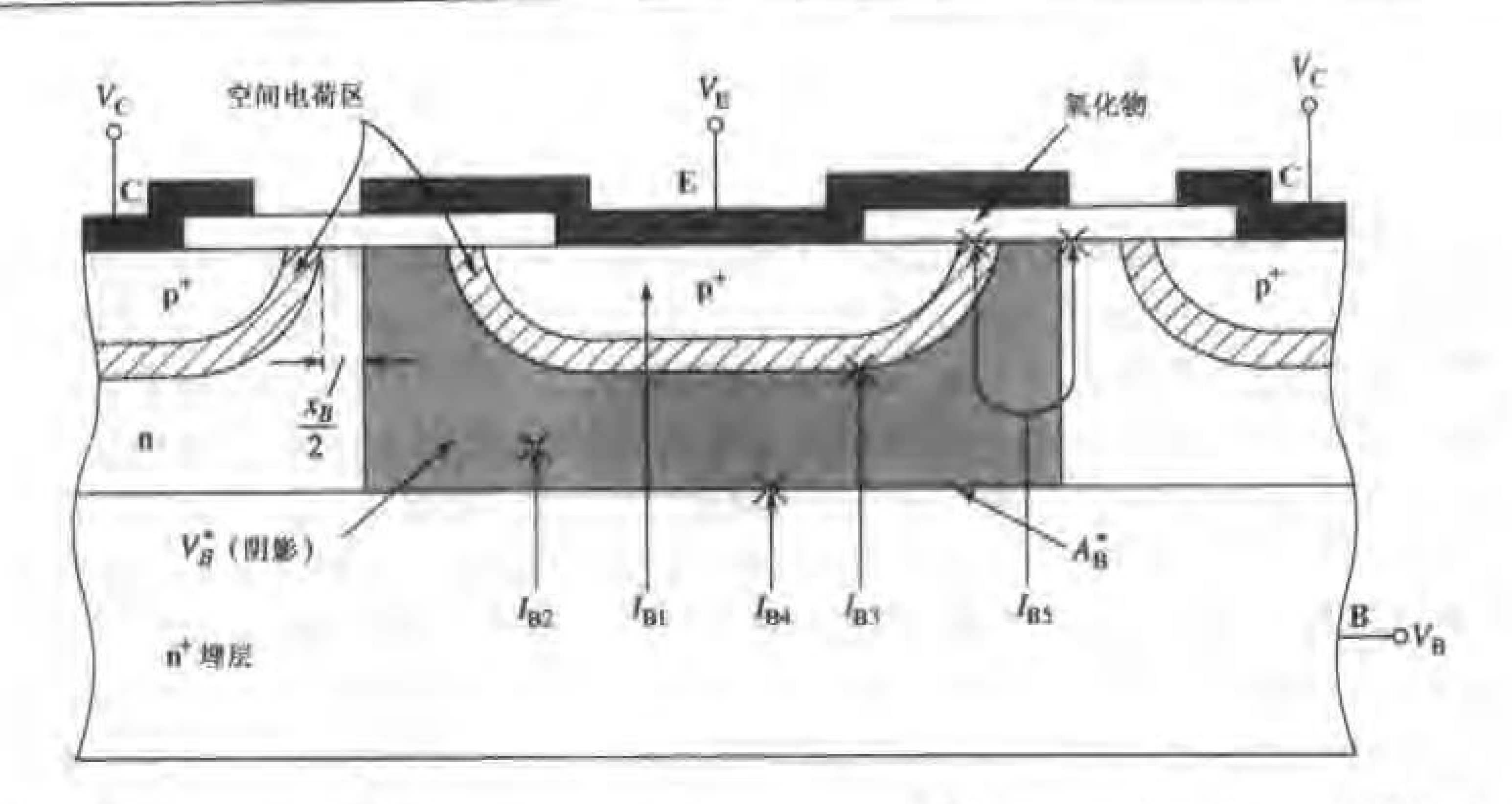


图7.37 横向pnp晶体管的基极电流示意图[20]。I_{B1}—注入到发射区的电子电流;I_{B2}—基区复合电流;I_{B3}—体内空间电荷区复合电流;I_{B4}—埋层复合以及被衬底收集的电流;I_{B5}—表面复合电流

式中 A_B^* 是 V_B^* 底部的面积(见图7.37)。横向pnp晶体管中没有真正具有复合速度 s_{nn^+} 的平面,然而,在计算流入埋层的空穴流时,复合速度是一个有效的参数。流入埋层的空穴可以在表面复合,也可以在埋层体内复合,或者被衬底和埋层之间的结收集。硅-二氧化硅界面复合电流也可以通过定义复合速度 s_{so} 来计算,表达式与等式(7.8.3)类似。

得到所有的基极电流成分的表达式后,如果已知 s_{nn^+},s_{so} 以及基区和空间电荷区的空穴寿命,我们就可以计算出基极电流。通常我们需要用特殊的测试结构来确定这些参数[20]。横向pnp晶体管的实验表明,一般 s_{so} 的值较小($s_{so}\sim 1-5$ cm/s)。因此,硅-二氧化硅界面的复合电流 I_{B5} 通常可以忽略。在有效的发射极-基极偏置范围内,其余的基极电流成分都是不可忽略的。埋层复合速度 s_{nn^+} 的变化范围相当大,已发表的范围为从10cm/s变化到2000cm/s,其原因是不同的掺杂浓度及其梯度,以及埋层的几何结构和制造工艺。当 $s_{nn^+}\gtrsim 100$ 时,纵向的空穴电流一般比较重要。

横向pnp晶体管的电流增益 β 小于npn晶体管,一般为20或者更小。经过仔细设计,β 值可以达到100。在小的微安量级,β 通常小于1。I_C 增加到大约100μA(发射区面积近似为 10^{-7}cm^2)时,β 开始上升。大电流时,β 随 I_C 的增加很快地下降。减小发射结空间电荷区复合电流可提高小电流区的 β,而大电流区 β 下降的原因是大注入效应。

图7.38是横向pnp晶体管横截面的显微照片(放大1500倍),利用了磨角和染色技术描绘出了集成电路中的扩散区。这种显微照片在评估IC工艺和查找缺陷时非常有用。



图7.38 横向pnp晶体管的横截面示意图(引自Signetics Corporation)

小结

为了得到更精确的双极晶体管模型,必须在第6章介绍的基本原理的基础上增加几项重要的物理效应。对于工作在正向放大区的晶体管,需要考虑的一个效应是Early效应,即集电结偏置对集电极电流的影响,可以用Early电压描述该效应。其他效应则限制了晶体管的偏置范围。双极晶体管的小电流工作范围通常受限于发射结空间电荷区复合,该复合导致注入基区的少子减少,使器件的电流增益(β_F)随直流偏置的下降而降低。

当晶体管工作在大电流区时,需要考虑其他几个重要的效应:首先,当发射结空间电荷区边界基区少子浓度增加到与掺杂浓度相比拟时,发射效率下降。其次,集电结耗尽区的空间电荷调制效应。当满足相应的大电流条件时,准中性基区的边界受到影响,一般是基区展宽。为分析Kirk效应,我们必须详细考虑晶体管的掺杂分布,求解相应的Poisson方程,该方程需要考虑由偏置电流决定的可动电荷的影响。第三,基区分布电阻的影响。由于有源基区的欧姆压降,导致发射结上离基极电极越远处的偏置电压越小。双极晶体管所有的大电流效应可能同时起作用,理解它们对于设计晶体管是非常重要的。而对于电路设计来说,精确的晶体管模型通常就足够了。

注入载流子通过准中性基区的输运方程(基区渡越时间)可以写成适用于任意基区杂质分布的形式。对该公式的理解及其推导有助于获得晶体管的电荷控制描述方法。考虑大注入对基区渡越时间的影响导致了Webster效应,即与过剩多子有关的内建电场缩短了基区少子渡越时间。

晶体管的电荷控制模型由一组线性微分方程组组成,它为电路设计提供了一套非常有用的晶体管描述方法。该模型在考虑器件的外部特性时比较有用;也就是说,电荷控制模型不会提供器件内部物理效应的准确信息。考虑小信号特性时,利用电荷控制模型可推导出小信号等效电路(混合Π模型)。在设计放大电路时,该等效电路非常有用。

电荷控制模型、EM模型和混合Π模型的参数之间的相互关系,对于理解这些模型,以及设计获得模型参数的试验非常有用。GP模型成功地描述了双极晶体管大部分重要的物理效应,适合于用计算机辅助模拟。该模型需要在基本的EM模型基础上增加几个参数。电荷控制模型可用于研究双极晶体管的高频性能,并定义了两个重要的参数:f_T是截止(特征或渡越)频率;f_{max}是最高振荡频率。异质结双极晶体管减轻了同质结晶体管中一些需要折衷处理的限制,使设计更为灵活。异质结双极晶体管还可形成电场来加速载流子通过基区。由于集成电路工艺的要求,经常需要使用横向pnp晶体管——晶体管发射区和集电区之间的区域平行于表面。决定横向pnp晶体管性能的物理效应通常与npn晶体管有所不同。

参考文献

1. J. M. Early, *Proc. IRE*, **40**, 1401 (1952)
2. F. A. Lindholm and D. J. Hamilton, *Proc. IEEE*, **59**, 1377 (1971).
3. C. T. Kirk, *IRE Trans. Electron Devices*, **ED-9**, 164 (1962)
4. H. C. Poon, H. K. Gummel, and D. L. Scharfetter, *IEEE Trans. Electron Devices*, **ED-16**, 455 (1969): Reprinted by permission.

5. P. E. GRAY, D. DEWITT, A. R. BOOTHROYD, and J. F. GIBBONS, *Physical Electronics and Models*. SEEC Volume II, Wiley, New York, 1964.
6. O. MANCK, H. H. HEIMEIER, and W. L. ENGL, *IEEE Trans. Electron Devices*, **ED-21**, 403 (1974).
7. W. M. WEBSTER, *Proc. IRE*, **42**, 914 (1954).
8. R. BEAUFOY and J. J. SPARKES, *Automat. Teleph. Elect. J.*, **13**, Reprint 112 (1957).
9. R. N. NOYCE et al, *Electronics*, July 21. 1969, p. 74.
10. P. E. GRAY and C. L. SEARLE, *Electronic Principles: Physics, Models and Circuits*, Wiley, New York, 1969.
11. A. SEDRA and K. SMITH, *Microelectronic Circuits*, Fourth Edition, Oxford Univ. Press, 1998.
12. H. D. BARBER, *Can. J. Phys.* **63**, 683 (1985).
13. L. H. CAMNITZ and N. MOLL in *Compound Semiconductor Transistors: Physics and Technology*, ed. S. TIWARI, IEEE Press, New York, 1993, p. 21.
14. D. J. ROULSTON and F. HEBERT, *IEEE Electron Device Lett.*, **EDL-7**, 461 (1986).
15. S.-Y CHIANG, D. PETTENGILL, and P. VANDE VOORDE, IEEE 1990 Bipolar Circuits and Technology Meeting, paper 8.1, p. 172.
16. R. G. MEYER and R. S. MULLER, *IEEE Trans. Electron Devices*, **ED-34**, 450 (1987).
17. S. E. LAUX and W. LEE, *IEEE Elect. Dev. Lett.*, **11**, 174 (1990).
18. C. C. MCANDREW, J. A SEITCHIK, D. F. BOWERS, M. DUNN, M. FOISY, I. GETREU, M. MCSWAIN, S. MOINIAN, J. PARKER, D. J. ROULSTON, M. SCHROTER, P. VAN WIJNEN, and L. F. WAGNER, *IEEE J. Solid-State Circuits*, **31**, 1476 (1996).
19. H. K. GUMMEL and H. C. POON, *Bell Syst. Tech. J.*, **49**, 827 (1970).
20. S. CHOU, *Solid-State Electron.*, **14**, 811 (1971).
21. A. S. GROVE, *Physics and Technology of Semiconductor Devices*, Wiley, New York, 1967, pp. 229, 240.
22. J. LOGAN, *Bell System Tech. J.*, **50**, 1105 (April 1971).
23. A. BAR-LEV, *Semiconductor and Electronic Devices*, Prentice-Hall International, Englewood Cliffs, N.J., 1984.
24. I. GETREU, *Modeling the Bipolar Transistor*, Tektronix, Inc., Beaverton, OR 97077, 1976.

习题

7.1 假设集电极电流完全由发射结和集电结之间的扩散电流组成，证明由等式 7.1.3 推导出的 Early 电压 V_A的关系式正确指出了原型晶体管的$\partial I_C/\partial V_{CB}$。

7.2 已知原型晶体管基区掺硼，掺杂浓度为 $10^{17}cm^{-3}$，集电区掺磷，掺杂浓度为 $10^{16}cm^{-3}$，中性基区宽度为 2.5μm，集电结为突变结，计算 $V_{CB}=0$ 时的 V_A。由于 Early 效应的影响，I_C的斜率是多少？

7.3* 比较习题 6.3 中两个晶体管的 Early 电压。假设两个器件的$\partial x_B/\partial V_{CB}$近似相等。

7.4† 已知 npn 晶体管准中性基区的掺杂浓度线性变化，发射结边界为 $10^{17}cm^{-3}$，集电结边界为 $10^{16}cm^{-3}$，基区宽度为 1μm，发射区和集电区的掺杂浓度均为 $N_d=10^{19}cm^{-3}$。

(a)绘出以下状态的少子分布 (i)热平衡状态，(ii)正向放大区，小注入；

(b)绘出基区"内建"电场；

(c)推导内建电场表达式，并求出其最大值；

(d)确定以下两个 Early 电压的近似比值 (i)正向放大区 Early 电压 V_A；(ii)反向放大区 Early 电压 V_B。

7.5 利用等式 7.1.4，定性讨论 V_A与集电结偏置的关系：(a)原型晶体管；(b)IC 中的放大晶体管。

7.6 产生电流集边的一个判据是横向基区上的电压降超过 kT/q。已知晶体管的 $\beta_F=50$，杂质分布如图 P7.6 所示，条形尺寸为 $Z_E=0.1cm$，$Y_E=2\times10^{-3}cm$（图 6.3），估算相应的集电极电流。可以证明，对于条形结构，基区分布电阻可以表示成

$$R_B=\frac{\bar{\rho}_B Y_E}{6x_B Z_E}$$

式中基区平均电阻率

$$\bar{\rho}_B\simeq\frac{A_E x_B}{\mu_p Q_{BO}}$$

μ_p为与 $N_a=Q_{BO}/qA_E x_B$相应的迁移率[21]。

7.7* 已知 pn 结的特性参数与习题 5.12 相同。如果将该 pn 结用做 pnp 晶体管的发射结，β_F最高能达到多少？

7.8* 习题 6.17 和习题 6.18 已经求出了晶体管的 β_F，并且其性能的退化是由于辐射。重新考虑该晶

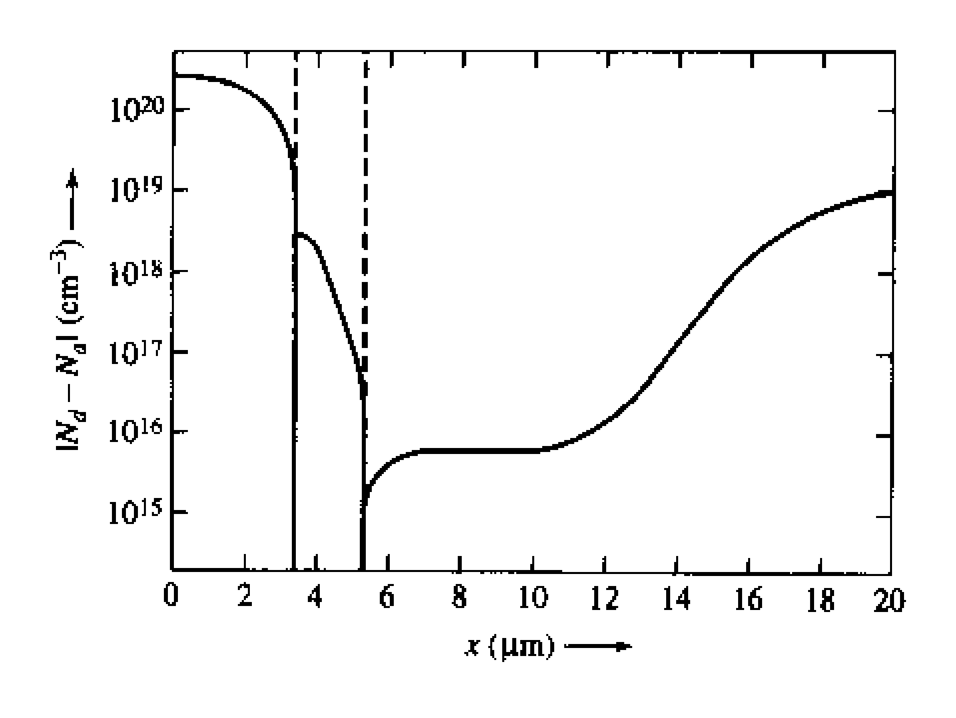


图 P7.6

体管,假设工作电压使等式 5.3.25 中 $J_t/J_r = 10$。

(a)计算这种情形下 β_F 的值。

(b)如果辐射导致的空间电荷区寿命的改变与基区少子相同,重新计算习题 6.18 的问题。

7.9* 推导 npn 原型晶体管的 β_F/β_{F0} 与集电极电流关系的表达式,并绘出曲线。要求考虑发射结边缘基区的大注入效应(见 7.2 节)引起的晶体管电流增益的下降。β_{F0} 为晶体管工作在中等电流区的电流增益。已知 $A_E = 5 \times 10^{-5}\mathrm{cm}^2$, $N_a = 5 \times 10^{16}\mathrm{cm}^{-3}$, $D_n = 20\mathrm{cm}^2\mathrm{s}^{-1}$, $x_B = 5\mu\mathrm{m}$,基极电流的主要成分是向发射区的反向注入。在这种条件下,基极电流仍保持正比于指数项 $\exp(V_{BE}/V_t)$。因此,当 $n(0) \sim V_{BE}$ 关系偏离中等电流区关系式时,β 开始下降。

7.10* 考虑 Kirk 效应(见 7.2 节)的极限情形:集电结空间电荷区移至 n^+ 埋层边缘。这种情形只需考虑负的空间电荷,正的空间电荷在高掺杂的埋层边缘变化非常陡峭。假设电子速度 v 为(a)$v = \mu\mathscr{E}$ 或(b)$v = v_l$,计算电场和空间电荷区宽度与电流的关系。这是固体中空间电荷限制电流的两种情形。

7.11 (a)说明通过类似图 7.10 所示的 $I_C \sim V_{BE}$ 曲线可得出等式(7.2.13)中的 R_B。特别是,如果用 I_{CA} 表示实际的集电极电流,I_{CI} 表示不考虑基极电阻的集电极电流(理想电流),证明

$$R_B = \frac{V_t\beta_F}{I_{CA}}\ln\left(\frac{I_{CI}}{I_{CA}}\right)$$

(b)利用图 7.10 的下列数据绘出 $R_B \sim I_{CA}$ 曲线。已知 $I_S = 3 \times 10^{-14}$ A, $V_t = 0.0252$V, $\beta_F = 100$(假设为常数)。

V_{BE}(V)	I_{CA}(mA)
0.70	11.25
0.72	22.4
0.75	56.2
0.80	200

7.12* 如果图 7.12 中的晶体管总电阻 $R = 150\Omega$(0.75□,方块电阻为 200Ω/□),外部电阻为 20Ω,利用图中的网络研究基极电阻导致的电流集边。假设每个子晶体管的饱和电流等于习题 7.11 给出的 I_S 的1/8,每个子晶体管的 $\beta_F = 100$。

(a)假设最里面的晶体管的电流分别为 1mA 或 10 mA,通过网络计算流过晶体管的总电流。

(b)利用习题 7.11 的结论计算 R_B 以及基极和发射极引线之间的总偏置。

(c)假设只考虑基区电阻引起的大电流效应，那么大电流时 R_B 的极限值为多少？

(d)证明电流非常小时，R_B 趋于外部电阻与 $11R/128$ 的和。

(为得出正确的结果，需要绘出大约四张与 V_{BE} 有关的图)

7.13　证明等式(7.3.3)。

7.14* 已知晶体管的基区掺杂为指数分布，基区(内建)电场为常数，基区的发射结边界 $x=0$ 和集电结边界 $x=x_B$ 之间的内建电压降为 κV_t，其中 V_t 为热电压 kT/q。利用等式(7.3.8)计算 τ_B。

(a)证明 $\nu=\kappa^2/(\kappa-1+e^{-\kappa})$，并且 $\kappa\to 0$ 时该式也正确。

(b)已知 $\kappa=20$，$x_B=0.5\mu m$，$D_n=20cm^2s^{-1}$，计算 τ_B，并解释为什么实际器件的 κ 值能达到该数值？

7.15† 考虑图7.8中Kirk效应对 τ_B 的影响，利用等式(7.3.8)绘出 τ_B 随集电极电流变化的半定量图。高频性能下降是Kirk效应的重要结果[1]

7.16　已知原型晶体管，发射效率由等式(6.2.20)给出，根据几何尺寸和少子寿命求出 τ_{BF}(由等式(7.4.3)引入)的表达式，并给出 τ_{BF} 接近基区电子寿命 τ_n 的条件。

7.17　证明pnp晶体管的电荷控制方程为

$$i_C=-\frac{Q_F}{\tau_F}+\frac{dQ_R}{dt}+Q_R\left(\frac{1}{\tau_R}+\frac{1}{\tau_{BR}}\right)+\frac{dQ_{VC}}{dt}$$

$$i_E=\frac{dQ_F}{dt}+Q_F\left(\frac{1}{\tau_F}+\frac{1}{\tau_{BF}}\right)-\frac{Q_R}{\tau_R}+\frac{dQ_{VE}}{dt}$$

$$i_B=-\frac{dQ_F}{dt}-\frac{Q_F}{\tau_{BF}}-\frac{dQ_R}{dt}-\frac{Q_R}{\tau_{BR}}-\frac{dQ_{VE}}{dt}-\frac{dQ_{VC}}{dt}$$

7.18　通过考虑空间电荷 Q_{VE} 和 Q_{VC} 的物理本质，说明方程组(7.4.11)以及上题(pnp晶体管)中这些量的微分项的符号是正确的。提示：考虑提供基极电流的电荷的符号和 Q_V 项的符号。参考图7.14会有一些帮助。

7.19　通过必要的步骤说明等式(7.4.10)是正确的。

7.20　在直流条件下分析图7.18(与等式(7.4.11)有关)中的电路，如果定义

$$\alpha_F=\frac{\tau_{BF}}{\tau_F+\tau_{BF}}$$

$$\alpha_R=\frac{\tau_{BR}}{\tau_R+\tau_{BR}}$$

和

$$I_{ES}=Q_{FO}\left(\frac{1}{\tau_F}+\frac{1}{\tau_{BF}}\right)$$

$$I_{CS}=Q_{RO}\left(\frac{1}{\tau_R}+\frac{1}{\tau_{BR}}\right)$$

证明等式(7.4.11)可简化成EM方程(等式(6.4.10))。

7.21† 证明等式(7.4.15)中的 τ_{SLOW} 可以写成

$$\tau_{SLOW}=\frac{(\beta_F+1)\tau_{BR}+(\beta_R+1)\tau_{BF}}{1+\beta_F+\beta_R}$$

这种表示 τ_{SLOW} 的方法在实践中经常用到。

7.22　(a)证明等式(7.4.2)中定义的 Q_{FO} 与 $n_{p0}(0)$ 的关系是线性的，并写出原型晶体管的 Q_{FO}。提示：考虑等式(7.1.1)和(7.4.1)。

(b)证明饱和晶体管的 V_{CE} 为

$$V_{CE}=\frac{kT}{q}\ln\left[\frac{Q_{RO}(Q_F+Q_{FO})}{Q_{FO}(Q_R+Q_{RO})}\right]$$

7.23* 已知晶体管的电荷控制参数为:$\tau_F=12\text{ns}$,$\beta_F=100$,$\tau_R=36\text{ns}$,$\beta_R=10$。

(a)如果晶体管工作在 $V_{CB}=0$ 的临界饱和区,集电极电流 $I_C=2\text{mA}$,计算正向存储电荷 Q_F。

(b)如果基极电流改为 $I_B=0.5\text{mA}$,I_C仍保持 2mA,计算 Q_F和 Q_R。

(c)比较(a)和(b)两种情形下的存储电荷。

7.24 用电容模型计算电荷随电压的变化时,不仅需要考虑与电荷相关的电容两端电压的变化关系,还需要确保电荷变化与电容器端电压的符号一致。说明用图 7.27 中的电容 ηC_D 模拟 Q_F 的变化时(等式(7.5.6)),该要求是满足的。

7.25 说明如果考虑 v_{BE}和电流源的符号,图 7.27 中的等效电路不仅适用于 npn 晶体管,还适用于 pnp 晶体管。

7.26 小信号等效电路的成立条件是基极-发射极电压的变化小于 V_t,试问推导等效电路的哪些步骤需要这个限制?

7.27 通过必要的步骤将图 7.25 中的等效电路简化至图 7.27 中的形式 (利用 $\delta\ll1$ 和 $\eta\ll1$)

7.28 已知晶体管工作在正向放大区,静态工作点为 I_E和 V_{BE},发射极电流增量为 dI_E。利用图 7.30 中的简化的混合 Π 等效电路(可忽略 C_{jC}),证明晶体管到达新的稳态所需要的时间为:

$$\tau_E\simeq\tau_F+(C_{jE}V_t/I_t)$$

这是影响f_T的延迟时间之一。

7.29* (a)求习题 6.1 和习题 6.3 中两个晶体管的混合 Π 等效电路。已知两个晶体管的 $I_C=2\text{mA}$,$\phi_i+V_{CB}=10\text{V}$,基区均匀掺杂,$\tau_n=100\text{ns}$。

(b)将两个晶体管用于小信号放大电路,请评价它们的性能。

7.30† 已知工作在正向放大区的 npn 晶体管,照射集电结空间电荷区,单位时间辐射产生的电子-空穴对为 r,已知 r 与时间为正弦函数关系

(a)简要指出产生载流子的流向。

(b)指出如何在低频混合 Π 等效电路(见图 7.28)中包含辐射效应(忽略 Early 效应)。

(c)如果 $v_{BE}=0$(基极-发射极交流短路),用(b)问中的电路计算辐射产生的电流 i_C。

(d)如果基极-发射极交流开路($i_B=0$),重复(c)。

7.31 如果定义

$$I_{AA}=I'_{ES}\left[\exp\left(\frac{qV_{BE}}{kT}\right)-1\right]$$

$$I_{BB}=I'_{CS}\left[\exp\left(\frac{qV_{BC}}{kT}\right)-1\right]$$

说明"传输型"EM 方程(等式(6.4.2)和(6.4.3))可以描述成图 P7.31 所示的对称的等效电路。其中 I'_{ES}和 I'_{CS}与原始型 EM 方程中的 I_{ES}和 I_{CS}不同。J. Logan [22]讨论了这种描述方式相对于传统的 EM 等效电路(见图 6.12)的优点。

7.32† 推导等式(7.7.2)。

7.33* 推导等式(7.7.9)中的 q_b的解。

7.34† 利用 7.7 节 GP 方程的推导和讨论,说明低偏置放大工作区下的 $\beta_F(I_C/I_B)$随电流的下降可以被预测,推导出低偏置区 $\beta_F\sim I_C$关系式,并说明参数 n_e可以通过上述特性获得。绘出合理的 $\beta_F\sim I_C$曲线,已知 I_C大于 0.9 mA 时 $\beta_F=100$,低电流区渐近线与中等电流区渐近线交点处的 $I_C=0.5\text{mA}$,请问如何从 β_F曲线获得 I_1的值?

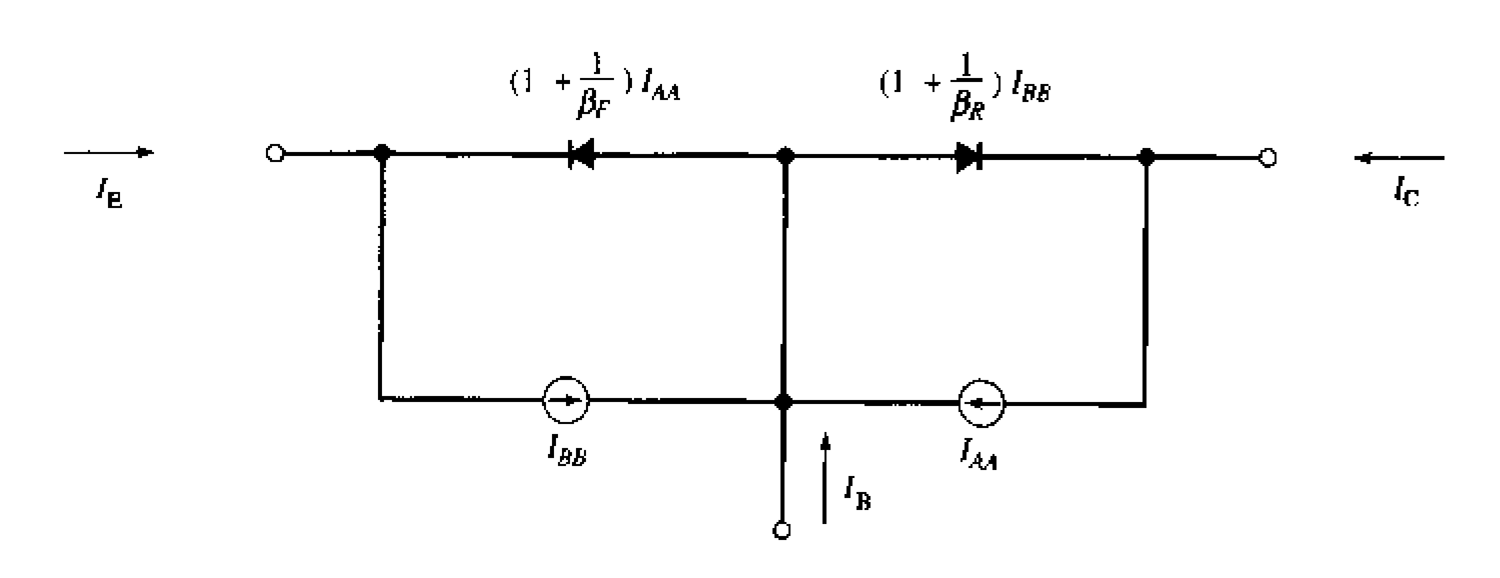


图 P7.31

7.35[+]　(a) GP 模型用两条渐近线形式描述集电极电流与基极-发射极偏置之间关系，请说明渐近线的交点为"膝点电流"$I_C = I_{KF}$（见图 7.32）.

(b)用 GP 模型证明在高电流区，β_F 与 I_C^{-1} 成正比。

7.36　使用合理的材料参数，计算基区延迟时间 τ_B：(1)基区均匀掺杂的硅双极晶体管；(2)基区杂质指数分布的硅双极晶体管；(3)基区带隙改变量为 0.2eV 的异质结双极晶体管(其他参数用硅的参数)。

第8章　金属-氧化物-半导体系统的性质

在第1章和第2章关于材料的电学特性和工艺技术的讨论中,我们知道硅因其特殊的性质已成为主流半导体材料。硅之所以能在众多的半导体材料中具有突出的优势,主要原因在于通过兼容的工艺,硅可被制成电学和机械性能良好的半导体(单晶硅)和绝缘体(无定型二氧化硅)。这一性能使得平面工艺成为可能,进而导致大规模集成电路得以可靠生产。氧化物-硅系统的性质是集成电路器件特性的基础。对这些性质的掌握和控制,导致在器件的设计和性能的提高方面取得了许多重要的进展。尽管这一领域的工作已经开展多年,但对氧化物-硅系统的研究仍在继续,对这一问题的理解也在不断加深。绘制能带图是研究氧化物-硅系统一个有效的出发点,尤其是当图中加入了第三种材料:即氧化物上方有金属层时,能带图的作用将更加明显。氧化物上方是用做电极的金属,这样就构成了三组分系统,即金属-氧化物-硅(MOS)系统。对MOS系统的分析有助于我们理解一些重要的集成电路结构:其中最重要的是金属-氧化物-硅场效应晶体管(MOSFET),有时也称做绝缘栅场效应晶体管(IGFET)[1]。

MOSFET是一种非常重要的器件,我们将在随后的两章中专门讨论。随着集成度的提高,超大规模集成电路已可在一个芯片上集成上千万个MOSFET,因而在超大规模集成电路中更多地采用MOSFET,而不是双极晶体管。本章关于MOS系统的讨论将帮助我们理解MOSFET的工作原理和物理机制。

本章主要讨论均匀掺杂的受主型(p型)硅的MOS系统,施主型(n型)硅的MOS系统的分析依此类推,有几道习题涉及该内容。在本章末的表8.3中有这两种系统的重要方程的总结。

下面将主要讨论二氧化硅-硅系统的性质。然而,随着MOS工艺的发展,高性能IC要求氧化层越来越薄,因此人们正在研究二氧化硅以外的其他绝缘材料,比如二氧化硅和氮化硅的结合或其他具有高电容率(或介电常数)的绝缘材料,如五氧化钽。尽管这些材料与硅的界面性能不如二氧化硅-硅系统,但仍具有很高的应用价值。很多关于氧化物-硅系统的讨论都适用于这些复杂的系统。

8.1　理想MOS结构

为获得金属-氧化物-硅系统的能带图,我们将利用前面的金属-半导体系统以及pn结的基本原理。首先我们认为在热平衡时,系统中各处的Fermi能级保持一致。因此,三组分系统,即金属-氧化物-硅中的Fermi能级为常数。暂且考虑理想情况,假设MOS系统的氧化层中及

---

[1] 尽管在很多情况下IGFET和MOSFET可以互换,但更精确地说,IGFET代表的范围更宽,并不只限于金属-氧化物-硅系统器件,而一般认为"MOSFET"就是指金属-氧化物-半导体场效应晶体管。

其界面处均无电荷。

8.1.1 热平衡能带图

电子从Fermi能级高（功函数小）的材料越过界面向Fermi能级低（功函数大）的材料迁移，使得不同材料中的Fermi能级达到一致。如第3章的讨论，真空能级是位置的连续函数。因此，知道了绝缘体和半导体的电子亲和势以及半导体和金属的功函数就可以确定惟一的能带图。图8.1所考虑的系统为铝（功函数=4.1V）、二氧化硅（电子亲和势~0.95V）和均匀掺杂p型硅（电子亲和势为4.05V，功函数为4.9V）。真空能级表示为E_0，材料接触前的各种能量也都在图中标出。

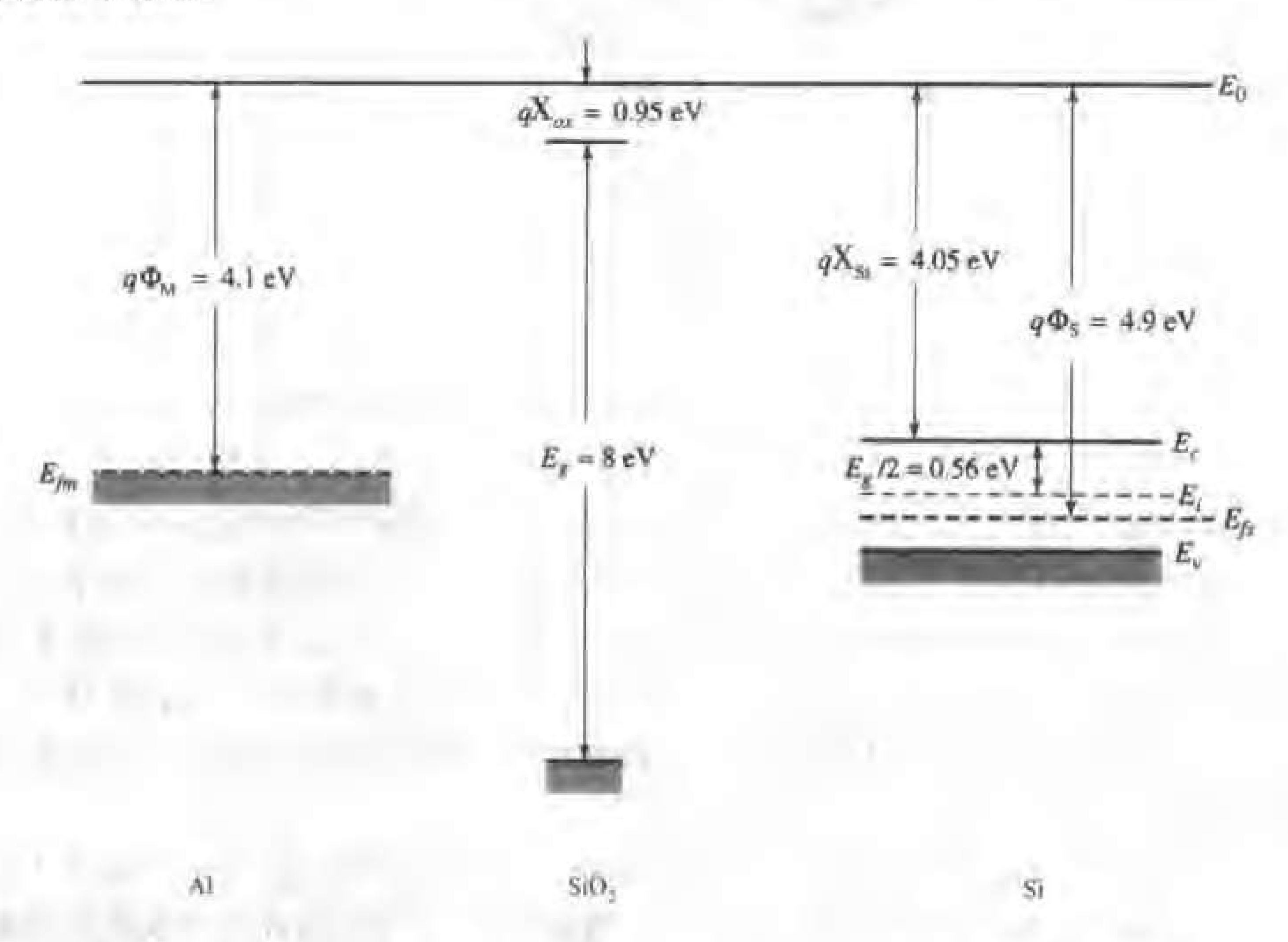


图8.1 组成MOS系统的三种材料在分离时的能带图：铝，热生长二氧化硅，$N_a\approx1.1\times10^{15}cm^{-3}$的p型硅（注意不同的表中给出的功函数和电子亲和势可能相差较大。这里给出的是最常用的值。）

当这些材料密切接触时，由于金属的功函数比硅低0.8V，所以负电荷从铝向硅转移，直到系统达到平衡。由于理想情况下绝缘层中没有可动电荷，不能够转移电荷，所以在绝缘层两侧将存储电荷，并产生电压降。在金属表面有一薄层正电荷（在良导体的理想情况下为面电荷），从硅表面延伸至体内的是带负电的受主层。与Fermi能量差对应的压降分别降落在氧化层和硅表面的空间电荷区上。

如MOS系统中是氧化层几乎完全绝缘的，这时的电荷迁移是很困难的。实际上，如果除了通过理想氧化层外，金属和硅之间没有其他的电荷通路，则这些材料将处于非平衡条件下

(即没有统一的 Fermi 能级)[2]。但是,几乎每一个 MOS 系统都有其他的电荷传输途径。例如,铝栅电极和硅衬底可以连在一起,或者在它们之间还有其他的欧姆导通途径。因此,可以假设金属和半导体之间存在热平衡状态。在这些假设条件下,由图 8.1 所示的材料形成的 MOS 系统其能带图如图 8.2 所示。

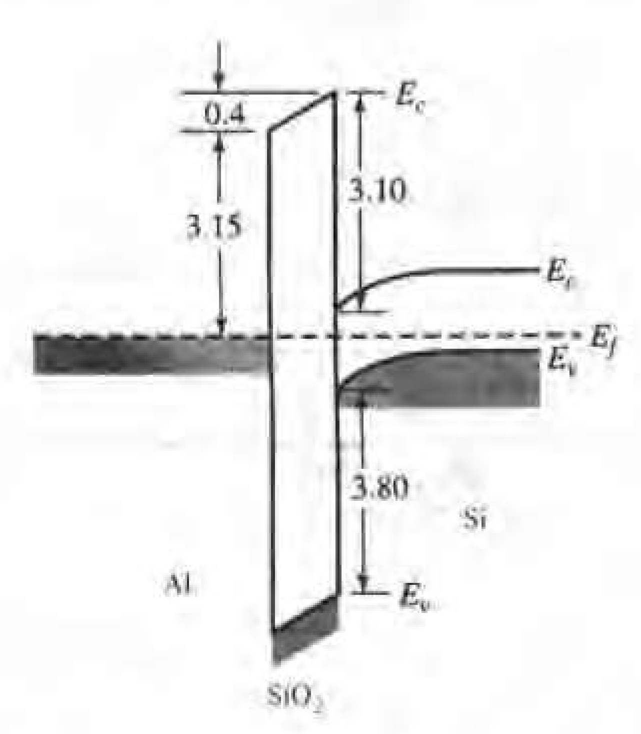


图 8.2　由图 8.1 所示的材料形成的理想 MOS 系统在热平衡时的能带图。假定 SiO_2 和 $Si\text{-}SiO_2$ 界面是理想的,且不带电荷

根据电荷流动,图 8.2 所示的能带图是这样得到的,p 型硅中的空穴流向欧姆接触(图中未画出),在欧姆接触处与电子复合。这些电子由 MOS 系统中的铝电极提供。由于电子的转移,靠近 SiO_2 的金属表面将出现一正电荷薄层(该层正电荷与硅表面存储的负电荷相等),所以铝不再保持电中性。这样,图 8.2 的理想 MOS 系统中,SiO_2 层两侧有 0.4V 的电压降。在实际 MOS 系统中,氧化层内或不同材料的界面上可能存在电荷,因此金属和半导体表面之间的能量差将有所改变,我们将在下面的例题中讨论。

图 8.2 的能带图为 p 型硅表面空穴耗尽的情况。我们可以将它与图 3.3 和图 3.4 中金属-半导体(金-n 型硅)系统的能带图作一比较。这两种情况下,硅表面的空间电荷均由离化的杂质原子组成,并且硅中的能带是弯曲的,这与求解 Poisson 方程得到的结论相一致。它们的区别在于氧化层的存在,图 8.2 中的氧化层使表面电荷分离,因而减小了表面电场。但是当硅表面耗尽时,这两种情况下硅体内的能带没有什么重大的区别。

在电学特性上这两种系统存在较大的差别。MOS 系统中,电子不能自由地从两个方向穿越氧化层,在能带图上表现为 SiO_2 层引入了较高的电子势垒。从图 8.2 可以看出,金属中 Fermi 能级处的允许态电子能量比二氧化硅导带允许态的能量低 3.15eV[3]。当电子从金属流向 SiO_2 时,可用高为 3.15eV 的势垒模型来表征金属-氧化物界面。同理,氧化物-硅界面也存在电子势垒,硅导带中电子的势垒高度为 3.10eV,价带电子为 4.20eV。通过测量的从金属和半导体向二氧化硅层注入电子所需的光子能量,证明了上述结论是正确的。

② 这一保持电荷的机制通常被用于不挥发存储器件,我们将在第 10 章讨论。

③ 如前所述,理论上是不能用能带的概念描述诸如二氧化硅这样的无定型材料的。但是就我们目前讨论的内容,这一概念是有用的。

例题　MOS 能带图

计算图 8.2 所示的 MOS 能带图中的二氧化硅层厚度是多少?

解:图 8.2 是图 8.1 所示材料在热平衡时的 MOS 系统能带图。金属和硅之间电势差(因为功函数的不同)等于 $\Phi_{MS}=4.9-4.1=0.8\text{V}$。由图 8.2 可知二氧化硅层两侧有 0.4V 的压降,所以降落在硅表面附近的压降也为 0.4V。

由于在 SiO_2 中不存在电荷,所以二氧化硅层中的电场$\mathscr{E}_{ox}$保持为常数,二氧化硅层两端的电压降 V_{ox} 可以简单地表示为$\mathscr{E}_{ox}\times x_{ox}$,其中 x_{ox} 表示二氧化硅层的厚度。因此,只要知道了$\mathscr{E}_{ox}$,就可以求出 x_{ox}。

由于假定氧化物-硅界面没有电荷,则垂直于界面的电位移矢量 D 是连续的,并且二氧化硅层中的电场与硅表面电场$\mathscr{E}_{s0}$有下述关系

$$\mathscr{E}_{ox}=\frac{\epsilon_s}{\epsilon_{ox}}\mathscr{E}_{s0}$$

硅的表面耗尽层的电荷密度为 qN_a,并且从 Si-SiO_2 界面向体内延伸,其厚度为 x_d。类似 3.2 节对 Schottky 二极管的讨论,这个区域的电场及电压与位置 x 有关。利用推导等式(3.2.2)和等式(3.2.3)类似的方法,可写出硅表面电场$\mathscr{E}_{s0}$以及耗尽层厚度 x_d的表达式

$$\mathscr{E}_{s0}=\frac{qN_ax_d}{\epsilon_s}$$

和

$$x_d=\left[\frac{2\Phi_s\epsilon_s}{qN_a}\right]^{1/2}$$

由图 8.1 知

$$(E_i-E_f)=(4.9-4.05-0.56)=0.29\text{eV}$$

在等式(1.1.27)中令 $p=N_a$ 可求出

$$p=N_a=n_i\exp[(E_i-E_f)/kT]=1.1\times10^{15}\ \text{cm}^{-3}$$

根据 $\phi_s=0.4\text{V}$,经计算得到 $x_d=685\text{nm}$,$\mathscr{E}_{s0}=1.17\times10^4\text{V cm}^{-1}$。因此

$$\mathscr{E}_{ox}=3.505\times10^4\ \text{V cm}^{-1}\quad 和\quad x_{ox}=\frac{V_{ox}}{\mathscr{E}_{ox}}=114\ \text{nm}$$

8.1.2　多晶硅栅和金属栅

由于硅表面的状况受金属电极的控制,所以金属层常常被称做栅,金属上的电压用 V_G 表示。铝一直是 MOS 工艺中最主要的金属栅材料,直到 20 世纪 70 年代后期引入了比铝金属性能更好的重掺杂多晶硅。由第 2 章知,生长了栅氧化层后,可立即用化学气相淀积(CVD)方法在栅氧化层上生长高纯硅,这样可以防止后续工艺对氧化层的污染。由于硅被淀积在无定型的二氧化硅上,所以形成的是由亚微米尺寸的晶粒组成的多晶薄膜。然后对多晶硅作 n 型或 p 型杂质的高掺杂,这将使多晶硅的导电性类似于金属。多晶硅的一个主要优点是能够承

受高温处理。多晶硅的自对准工艺可以精确定位 MOS 晶体管的源漏区，并减小覆盖电容。

20 世纪 90 年代 MOS 工艺进入到亚微米阶段，多晶硅栅的缺点也开始表现出来，主要是它的电阻比金属的高，即使是在高掺杂的情况下。人们正在研究新的栅材料，尤其是难熔金属，如钨。然而，采用这些新材料会产生污染，并且与其他材料在工艺不兼容，所以还没有进入实用阶段。但人们仍在尝试寻找其他可替代多晶硅的栅材料。

8.1.3　平带电压

在热平衡状态下的理想 MOS 系统中，金属和半导体构成了电容的两个极板。这个电容器上的电压与金属和半导体之间的功函数差相对应。在金属和硅之间加偏置，可使系统偏离热平衡状态，且使电容两端存储的电荷也发生相应的改变。考虑图 8.2 中的情况，金属相对于硅加负偏压时，将减小电容上的内建电压，并使电容极板上存储的电荷减小到低于平衡值。可以把负栅压的作用看做是将带正电的空穴拉向半导体表面而中和了一部分带负电的电离受主。

当外加电压恰好补偿了金属和半导体之间的功函数差时，MOS 电容两端存储的电荷减少至零，并且氧化层和半导体中的电场也消失了。在这种情况下，硅中的能带是水平的，或者说表面区与体内中性区(见图 8.3)的能带是一致的。由于外加电压的影响，使硅中能带处于水平时的外加电压称做平带电压，通常用 V_{FB} 表示。平带电压会随硅的掺杂浓度以及 MOS 系统金属栅的性质而变化。需要注意的是，在平带条件下 MOS 系统并不处于热平衡状态；因此，金属和硅中的 Fermi 能级并不相同(见图 8.3)。理想 MOS 系统的平带电压等于金属和硅的功函数差。

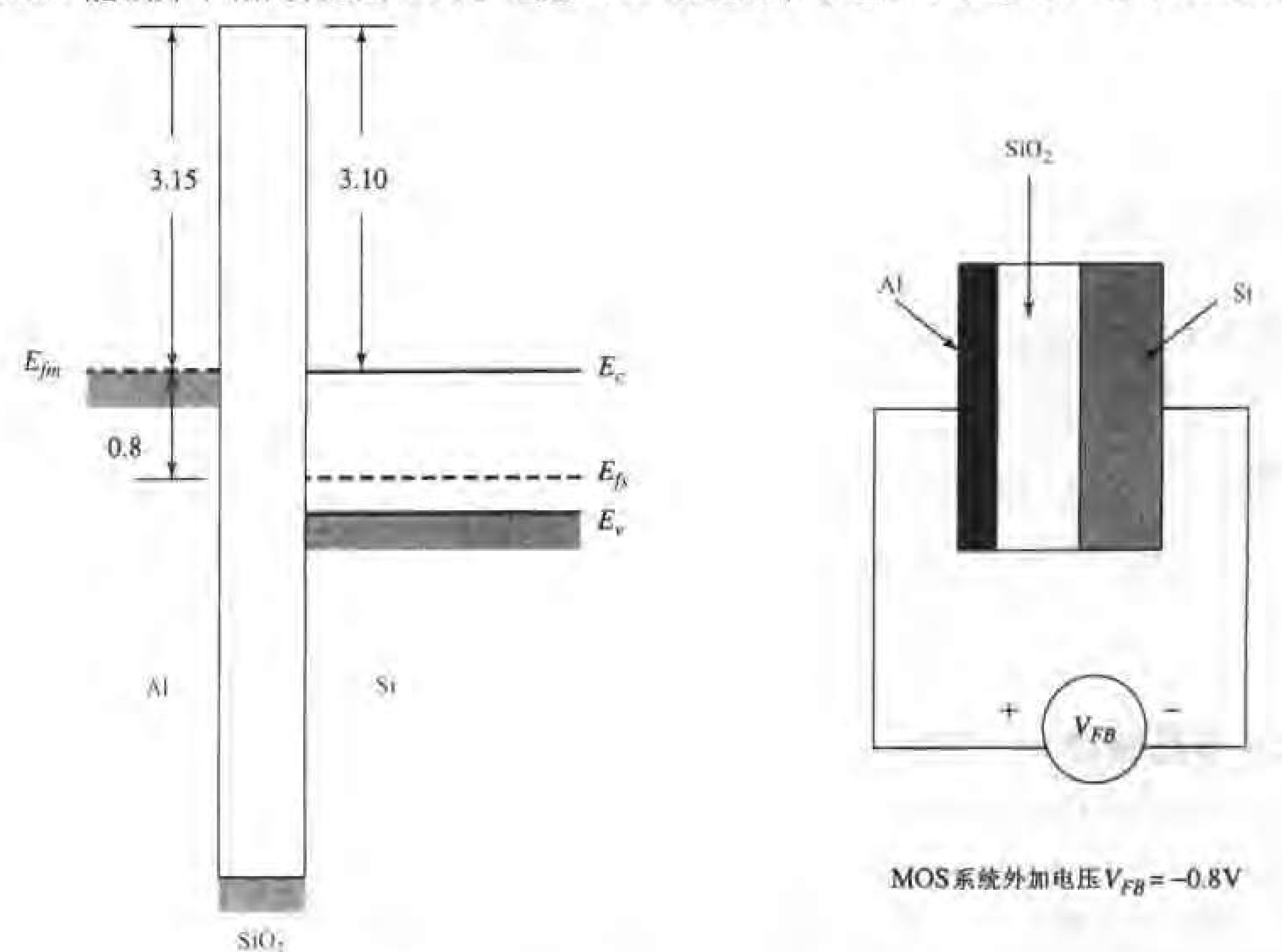


图 8.3　图 8.2 所示的 MOS 系统在平带情况时的能带图。金属与硅之间的外加偏压等于平带电压 $V_{FB}=-0.8$V，这时系统并不处于热平衡状态

$$V_{FB} = \Phi_M - \Phi_S = \Phi_{MS} \tag{8.1.1}$$

理想 MOS 系统中，不同栅材料的平带电压见表 8.1。

表 8.1 常用的栅材料和 $N_a = 1.1 \times 10^{15}\mathrm{cm}^{-3}$ 的 p 型硅的功函数（Φ_M和 Φ_S）和平带电压

栅材料参数	铝	n^+多晶硅	p^+多晶体	钨
Φ_M(V)	4.1	4.05	5.17	4.61
Φ_S(V)	4.9	4.9	4.9	4.9
V_{FB}(V)	−0.8	−0.85	0.27	−0.29

8.2 理想 MOS 结构的分析

定性描述

现在继续讨论图 8.2 所示的 MOS 系统。如将硅接地，金属栅加负电压，并且其数值大于 V_{FB}的绝对值，那么带正电的空穴将被吸引到硅的表面，MOS 电容开始存储正电荷，在硅表面形成空穴的积累，此时硅表面的空穴浓度远大于 N_a。这种情况称为表面积累（简称为积累），表面空穴积累的区域就是积累层。

表面积累层是由自由载流子形成的空间电荷层，可以用 3.4 节讨论的 Schottky 欧姆接触来描述其空间分布。由等式(3.4.2)中可看出，有一半的自由载流子分布在$\sqrt{2}L_D$中（L_D为 Debye 长度），因此表面积累层总的范围仅是几个 Debye 长度。为估计这个量级，现考虑表面积累的 p 型硅，其掺杂浓度 $N_a = 10^{15}\mathrm{cm}^{-3}$，表面载流子浓度是体内的 10 倍。此时表面 Debye 长度大约为 40nm（由等式(3.4.3)计算），这个值远大于目前多数 MOSFET 的栅氧化层厚度。图 8.4(a)和图 8.4(b)分别给出了 MOS 结构表面积累时的能带图和电荷分布。

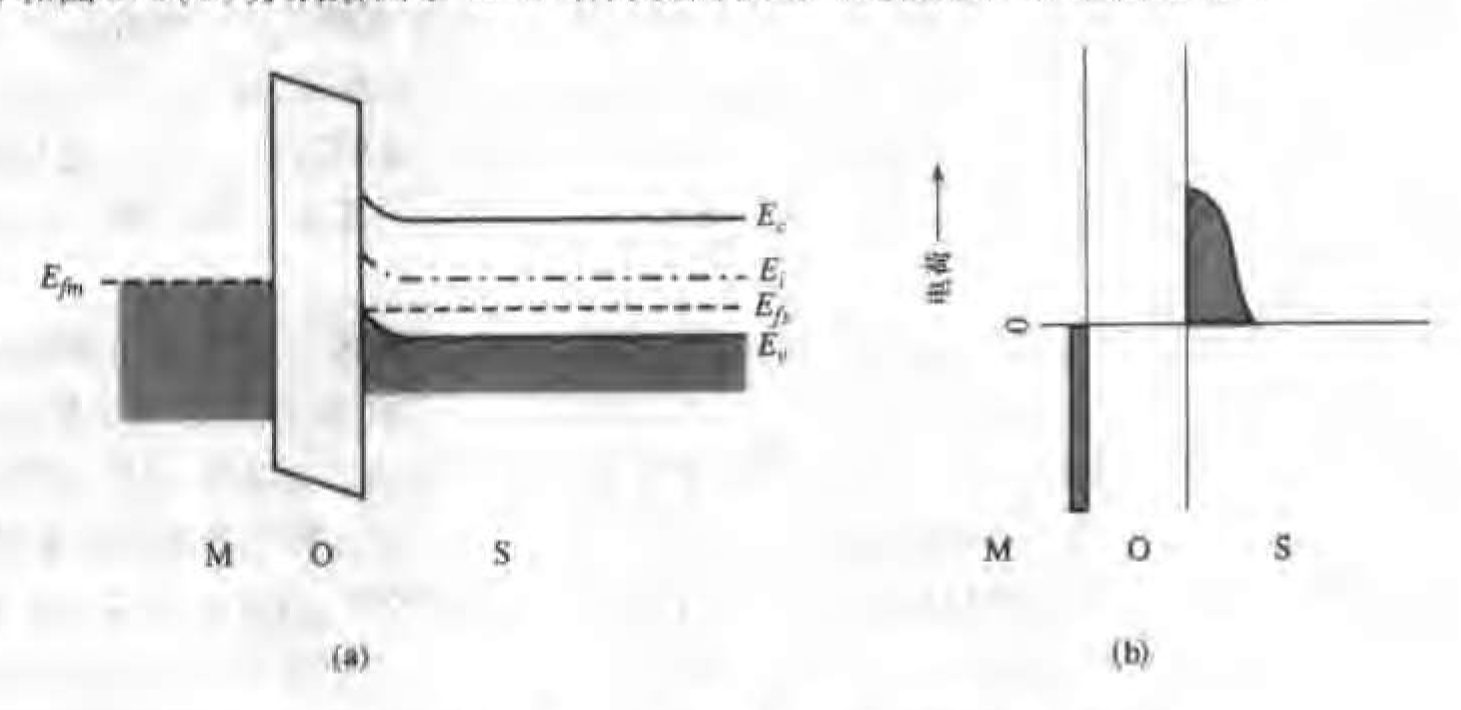


图 8.4 MOS 结构表面积累时的(a)能带图，(b)电荷分布

现在讨论其他偏置情况。从图 8.2 的 MOS 系统已知，零偏压时硅表面存储负电荷，金属表面存储正电荷。硅表面的负电荷是由空穴耗尽后的电离受主组成的。这些电荷的分布与金属和硅之间的内建正电压一致。当外加正电压时，硅表面将更加耗尽，更多的空穴被排斥而留

下更多的未补偿的受主离子。相应地,金属上的正电荷也会增加。由于硅中可动电荷远离表面,这种情况称为表面耗尽。图 8.2 所示的 MOS 系统在耗尽偏置下的能带图及电荷分布如图 8.5 所示。耗尽偏置时 MOS 结构的能带图与反偏 Schottky 势垒二极管非常相似,比较图 8.5(a)和图 3.5(b)可看出其相似性。(不同的是图 3.5 中为 n 型硅而图 8.5 中为 p 型硅。)

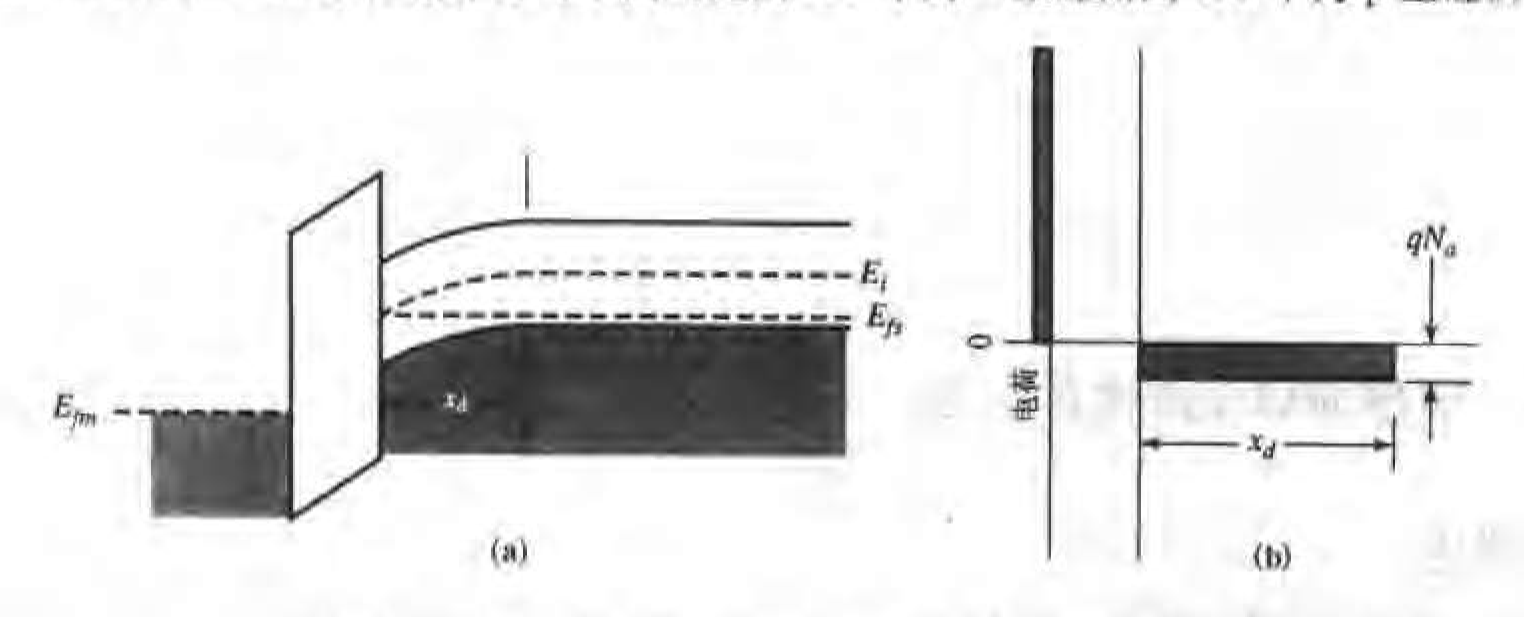


图 8.5 MOS 系统表面耗尽时的(a)能带图,(b)电荷分布

如继续增加 MOS 金属上的正电压,将会发生不同于金属-半导体二极管的现象。在 MOS 系统中,随着金属上电压的增加,硅表面的电场也将增加,表面处的能带发生显著弯曲。表面区域的多数载流子被耗尽,所以产生将大于复合(见等式 5.2.9)。这些产生的电子-空穴对被电场分离,空穴被扫入体内,而电子则移向 $Si\text{-}SiO_2$ 界面,但受到导带势垒的阻挡。假如金属上的电压变化足够缓慢,则这个产生过程就可以使硅表面的自由载流子浓度与体内的浓度达到平衡,这样,从硅的体内到二氧化硅界面就可以画出统一的 Fermi 能级,并且可以用 Fermi-Dirac 统计来计算硅中的载流子浓度。事实上,本节中所有的能带图都采用了这一假设(我们将在后面考虑其他情况)。假如外加电压变化导致能带弯曲时,硅中的 Fermi 能级仍保持不变,那么在足够高的电压下,表面本征 Fermi 能级 E_i 将与体内 Fermi 能级相交。在靠近 $Si\text{-}SiO_2$ 界面的硅体内,Fermi 能级离导带边比离价带边更近。就载流子浓度而言,这意味着外加电压产生了一个反型层,因为表面处的电子浓度比空穴高,尽管硅是受主型掺杂。此时,金属和硅之间所加的电压在表面附近感生出一个 pn 结。图 8.6(a)中给出了图 8.2 所示的 MOS 系统在反型时的能带图。

当表面处的 E_i 略低于 E_f 时,反型层中的电子浓度较低(为 n_i 量级),此时 MOS 系统处于弱反型区。当表面处的($E_c - E_f$)低于体内的($E_f - E_v$)时,反型层中的电子浓度将远大于体内的空穴浓度,此时系统处于强反型区。我们将表面电子浓度等于体内掺杂质浓度作为这两种状况的分界。虽然这样划分强反型区和弱反型区有些不确定性,但是这种区分方法很方便。

在耗尽和反型偏置下,外加电压通过排斥空穴产生表面耗尽区,同时吸引电子形成反型层,从而使半导体中的负电荷增加。图 8.6(b)表示了图 8.2 所示 MOS 系统在反型时的电荷分布。反型层中的自由电子电荷密度,称为 Q_n($C\ cm^{-2}$),这些电荷由于表面电场的作用而位于接近表面处。受主电荷则分布在从表面延伸至体内的整个耗尽区中。与第 3、4 章类似,我们用 Q_d 表示耗尽区电荷密度。$Q_n + Q_d$ 的和用 Q_s 表示,称为硅的空间电荷密度。

当硅进入强反型后,因为表面处的自由电子浓度与表面势成指数关系,所以在反型层形成后随栅压增加表面势变化很小。由栅压的增加感应出的电荷全部是位于 $Si\text{-}SiO_2$ 界面下方的

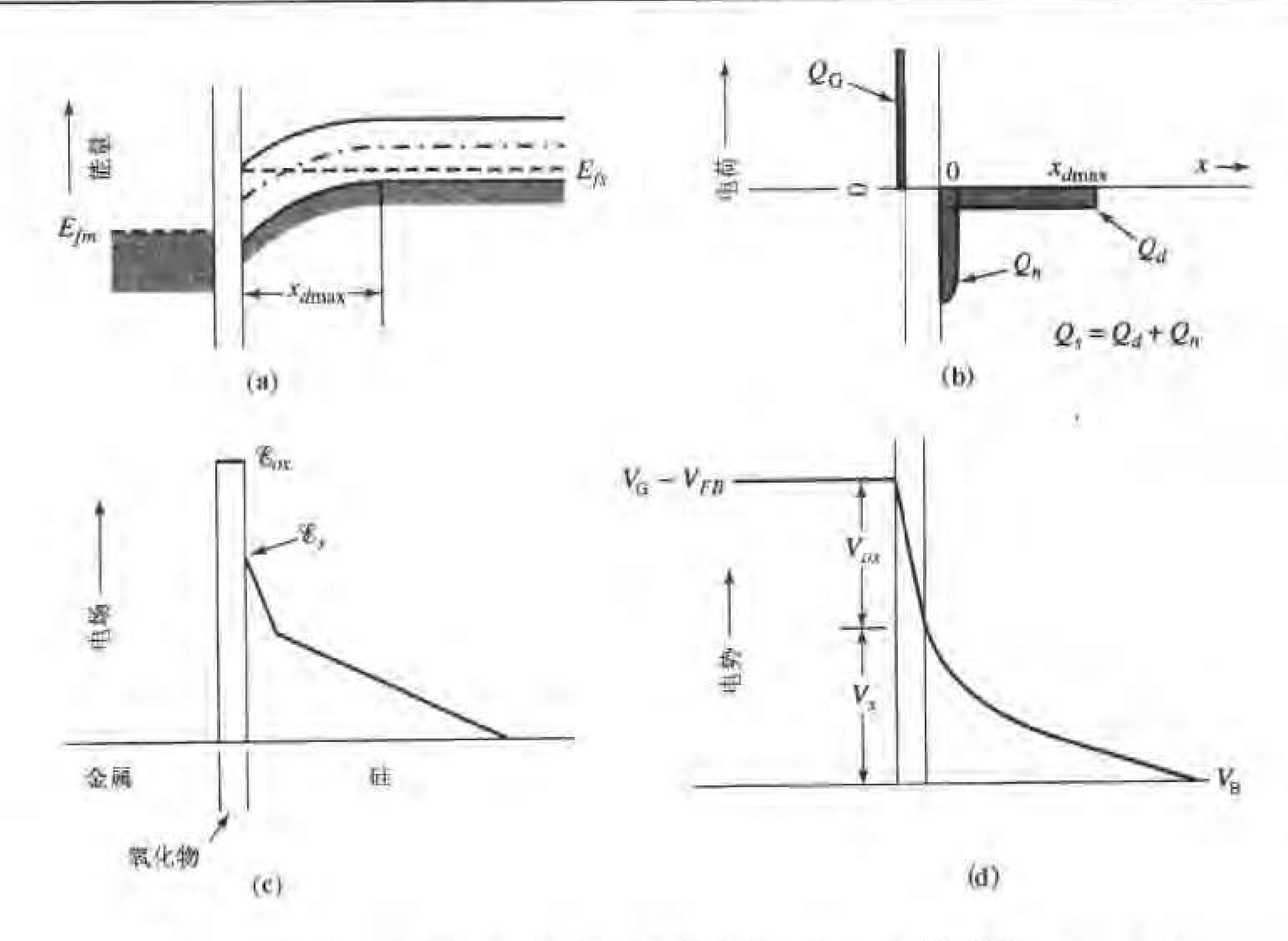


图 8.6　p 型硅衬底的 MOS 系统在反型时的（a）能带图，（b）空间电荷分布，（c）电场分布，（d）电势分布

自由电子。因此表面反型之后，硅中的耗尽区上的总压降和耗尽层宽度都相对不变。耗尽层最大宽度通常用图 8.6(b) 中的 x_{dmax} 来表示。图 8.6(c) 和图 8.6(d) 分别给出了反型 MOS 结构中的电场分布和电势分布。由于二氧化硅和硅的介电常数不同，所以界面处的电场不连续，从 $\mathscr{E}_{ox}$ 降低到 $\mathscr{E}_s$。MOS 结构两端的总压降（$V_G - V_B - V_{FB}$）由氧化层上的压降和硅中空间电荷区上的压降两部分组成。8.3 节中将推导出这两个电压的表达式。

上述对表面电荷、电压和电场的定性讨论说明了加在金属上的栅压对硅表面特性的重要作用。尽管这个系统从本质上讲只是一个电容器，但是硅表面不同的电荷分布形式会引起明显不同的电学特性。例如，当偏置在积累区时，硅表面表现出良好的导电性，并与体内的自由载流子相连；当偏置在耗尽区时，硅表面呈现出高阻状态；当偏置在反型区时，硅表面也表现出良好的导电性，但与硅体内相隔离。表 8.2 总结了 p 型硅 MOS 结构的各种表面状况。

表 8.2　p 型硅 MOS 结构的各种表面状态

$(V_G - V_{FB})$	ϕ_s	表面电荷状态	表面载流子浓度
负	负 $\lvert\phi_s\rvert > \lvert\phi_p\rvert$	积累	$p_s > N_a$
0	负 $\phi_s = \phi_p$	中性（平带）	$p_s = N_a$
正（小）	负 $\lvert\phi_s\rvert < \lvert\phi_p\rvert$	耗尽	$n_i < p_s < N_a$
正（大）	0	本征	$p_s = n_s = n_i$
正（大）	正 $\lvert\phi_s\rvert < \lvert\phi_p\rvert$	弱反型	$n_i < n_s < N_a$
正（大）	正 $\phi_s = -\phi_p$	强反型开始	$n_s = N_a$
正（大）	正 $\lvert\phi_s\rvert > \lvert\phi_p\rvert$	强反型	$n_s > N_a$

8.3 MOS 电学特性

8.3.1 硅衬底中的电荷模型

前两节定性介绍了由栅和衬底之间的外加电压在硅表面感应的电荷分布。为了后面两章能利用这一理论研究 MOSFET,还需要更多的定量分析,尤其是对表面反型状态。下面将介绍栅和硅体内电荷分布的简化模型,这个模型不需要精确描述载流子的浓度分布。因为开发一个能满足我们需要的 MOS 系统模型不需要非常精确的表达式。这一模型的完整求解发表在参考文献[1,2]中。用计算机模拟某些偏置区时也会用到这个模型。

8.3.2 热平衡状态

作为一级近似,考虑硅表面区与体内处于热平衡状态。由热平衡得到的结论一般也适用于非平衡状态。由等式(4.1.2)定义硅中的电势

$$\phi(x) = \frac{1}{q}[E_f - E_i(x)] \tag{8.3.1}$$

因为考虑的是热平衡状态,所以 E_f保持水平而 E_i随位置变化。图 8.7 所示的是 p 型材料,E_f低于 E_i,所以硅体内中性区的电势 ϕ_p为负。表面势 ϕ_s为

$$\phi(0) = \phi_s = \frac{1}{q}[E_f - E_i(0)] \tag{8.3.2}$$

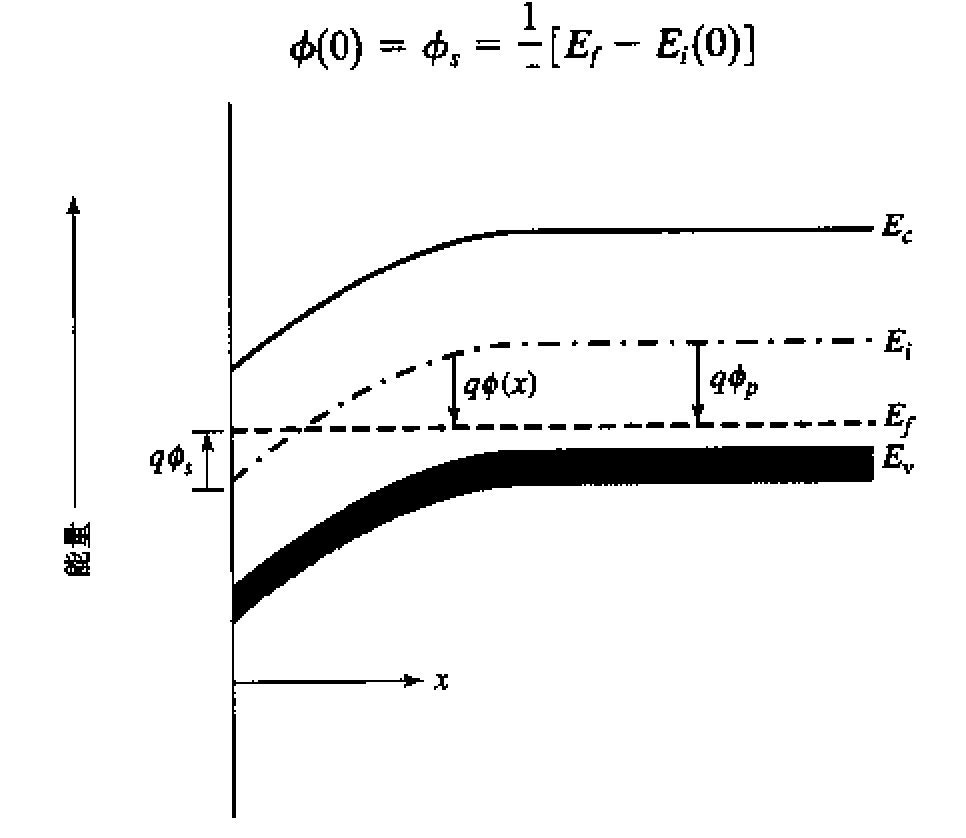


图 8.7 MOS 系统中硅表面附近的能带图,电势分布由等式(8.3.1)定义,图中的表面势(ϕ_s)为正

根据等式(1.1.26)和等式(1.1.27),得到载流子浓度与电势 $\phi(x)$的关系

$$p = n_i \exp\left(-\frac{q\phi}{kT}\right) \quad 和 \quad n = n_i \exp\left(\frac{q\phi}{kT}\right) \tag{8.3.3}$$

从这些等式以及 ϕ_p 和 ϕ_s 的定义，可用硅表面耗尽区上的电压降（$\phi_s - \phi_p$）来表示表面自由载流子浓度 n_s 和 p_s

$$p_s = N_a \exp\left[\frac{q(\phi_p - \phi_s)}{kT}\right] \qquad (8.3.4)$$

$$n_s = \frac{n_i^2}{N_a} \exp\left[\frac{q(\phi_s - \phi_p)}{kT}\right]$$

正如上一节所讨论的，当 MOS 系统处于积累区或反型区时，体内的自由载流子聚积在 Si-SiO_2 的界面处。根据热力学统计，等式(8.3.4)表明电荷浓度与该处的能带弯曲量呈指数关系。前面已经提到，当表面强反型或积累时，表面电势 ϕ_s 保持相对恒定，因为此时势 ϕ_s 只有很小的变化，n_s 和 p_s 就迅速改变。硅中的空间电荷密度 Q_s 的精确解中包括自由载流子电荷和耗尽区电荷，图 8.8 给出了 Q_s 的精确值与衬底到表面的总能带弯曲量，即 ϕ_{bs}（$\phi_{bs} = |\phi_s - \phi_p|$）的关系曲线（实线）。由于自由载流子浓度与所在位置的能带弯曲量为指数关系，所以反型层和积累层的物理厚度（例如，定义为包含 90% 这些电荷的区域）都仅有 ~10nm 量级。

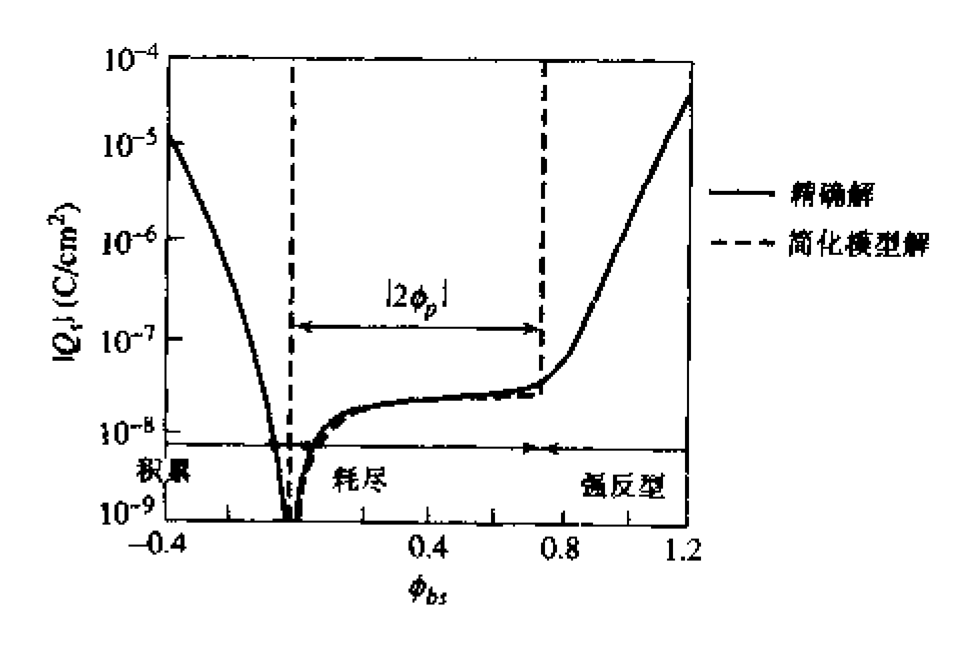


图 8.8　硅衬底中总的电荷 Q_s 与衬底到表面的总能带弯曲量 ϕ_{bs} 的关系曲线

简化模型假设积累开始于平带状态（衬底中总的能带弯曲量为零），反型开始于表面能带弯曲量 ϕ_{bs} 达到 $2\phi_p$ 时。此外，我们还假设反型层和积累层无限薄，自由载流子在硅表面形成面电荷层，并且反型层和积累层中没有能带弯曲。当表面势处于积累和反型之间时，硅处于耗尽状态，假设此时存在于耗尽区的空间电荷就是未被补偿的离化杂质。[4]

根据这个简化模型，利用类似第 3 章研究 Schottky 势垒的方法（等式(3.2.3)），可以利用耗尽近似，将表面势 ϕ_s 与耗尽层宽度 x_d 联系起来

$$x_d = \sqrt{\frac{2\epsilon_s|\phi_s - \phi_p|}{qN_a}} \qquad (8.3.5)$$

耗尽层电荷密度 Q_d 可以由下式给出

[4] 译者注：这里定义的耗尽区包括 8.2 节提到的弱反型区；所定义的反型区也称为强反型区。在有些论著中，当表面处的 E_i 略低于 E_f 时，称 MOS 系统处于反型区；当表面处的（$E_c - E_f$）低于体内的（$E_f - E_i$）时，称 MOS 系统处于强反型区。

$$Q_d = -qN_a x_d$$

当表面开始强反型时($\phi_s = |2\phi_p|$时)达到最大耗尽层宽度,因此有

$$x_{d\max} = \sqrt{\frac{4\epsilon_s|\phi_p|}{qN_a}} \tag{8.3.6}$$

和

$$Q_{d\max} = -qN_a x_{d\max} = -\sqrt{4\epsilon_s qN_a|\phi_p|} \tag{8.3.7}$$

图 8.8 中给出了简化模型求出的电荷密度 Q_s 与硅表面势 ϕ_s 的关系曲线(虚线)。

例题　Si-SiO$_2$界面附近的电势分布

在耗尽近似下推导理想 MOS 电容中的电势分布,并用表面电势 ϕ_s 和耗尽层宽度 x_d 来表示。已知硅为 p 型,取硅体内的电势为零, Si-SiO$_2$界面处 $x=0$。

解:利用耗尽近似及 Poisson 方程可知,硅体内 $0 \leqslant x \leqslant x_d$ 处的空间电荷和电场梯度为负值且保持恒定。如果表面电场为$\mathscr{E}_s$,那么电场与 x 的关系可写成

$$\mathscr{E}(x) = \mathscr{E}_s(1 - x/x_d) \qquad 0 < x < x_d$$

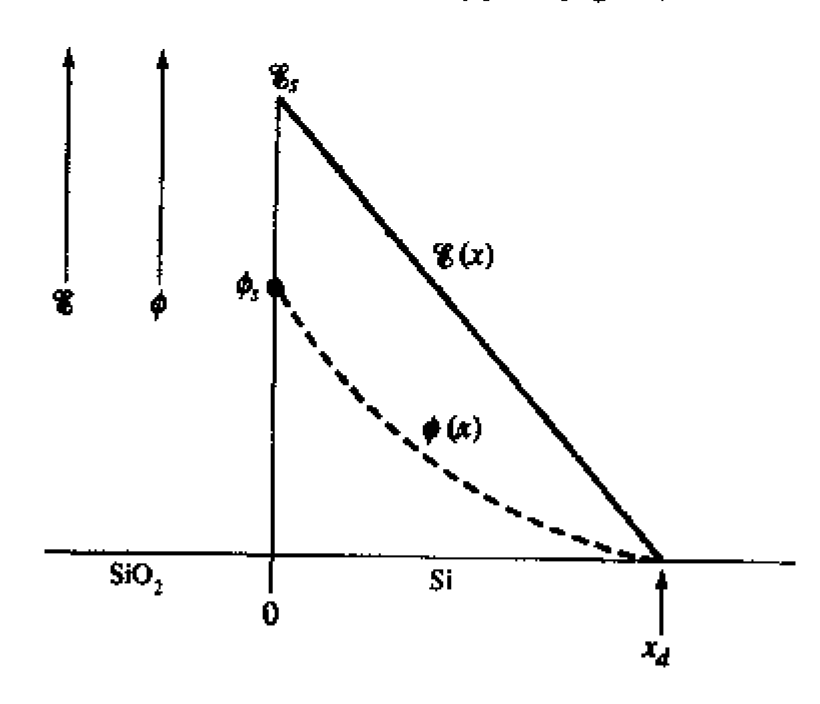
和

$$\begin{aligned}\phi(x) &= \phi_s - \int_0^x \mathscr{E}\,\mathrm{d}x \\ &= \phi_s - \mathscr{E}_s x + \mathscr{E}_s x^2/2x_d \\ &= \frac{1}{2}\mathscr{E}_s x_d - \mathscr{E}_s x + \mathscr{E}_s x^2/2x_d \\ &= \frac{\mathscr{E}_s}{2x_d}(x_d - x)^2\end{aligned}$$

所以电势与 x 的关系为

$$\phi(x) = \phi_s\left(1 - \frac{x}{x_d}\right)^2 \qquad 0 < x < x_d$$

表面势 ϕ_s 等于$\mathscr{E}_s \sim x$ 曲线下的面积。因此 $\phi_s = \mathscr{E}_s x_d/2$。



8.3.3 非平衡状态

当 MOS 系统强反型时，在硅的表面和体内形成了一个 pn 结。如图 8.9 所示，假如在该系统附近有一个 n 型扩散区与表面反型层相连接时，就可以在 pn 结上加偏置电压。此时，所加的电压使硅体内进入非平衡状态，并且在表面反型层和体内之间产生电流。但实际上，由于这个 pn 结是反偏的，产生电流很小。

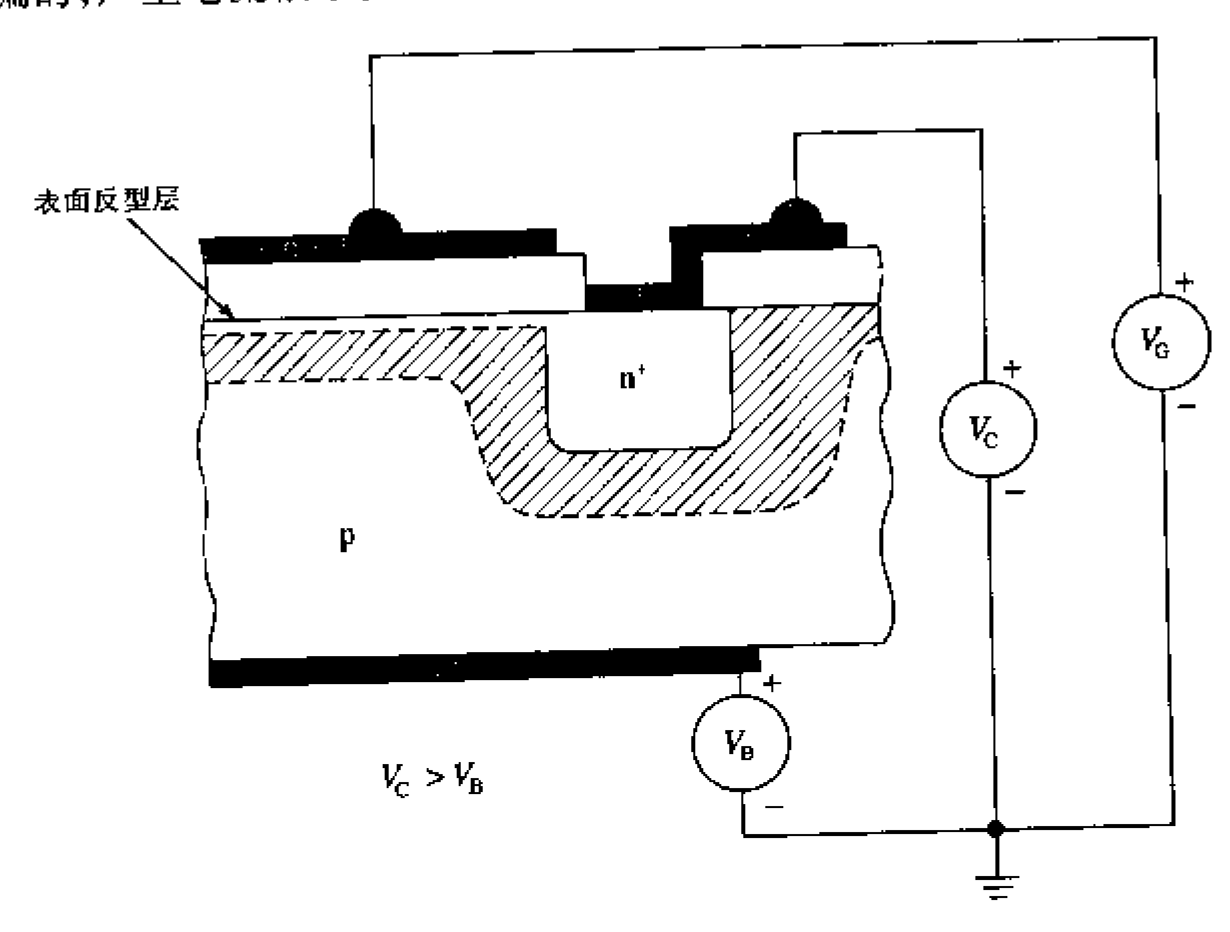


图 8.9　邻近 MOS 电容的扩散区可用来在体硅和 $Si\text{-}SiO_2$ 界面处的反型层之间加偏置电压。阴影部分表示硅体内耗尽的空间电荷区

可以用两个准 Fermi 能级(等式(1.1.28)和等式(1.1.29))描述偏置条件下反型层表面的能带图，一个是 p 型区的准 Fermi 能级，另一个是 n 型区的准 Fermi 能级。反偏 pn 结时这两个准 Fermi 能级之差等于外加偏压(见第 4 章)。如图 8.10 所示的反型层(或沟道)和 p 型衬底(或体)之间的外加反偏电压为($V_C - V_B$)。

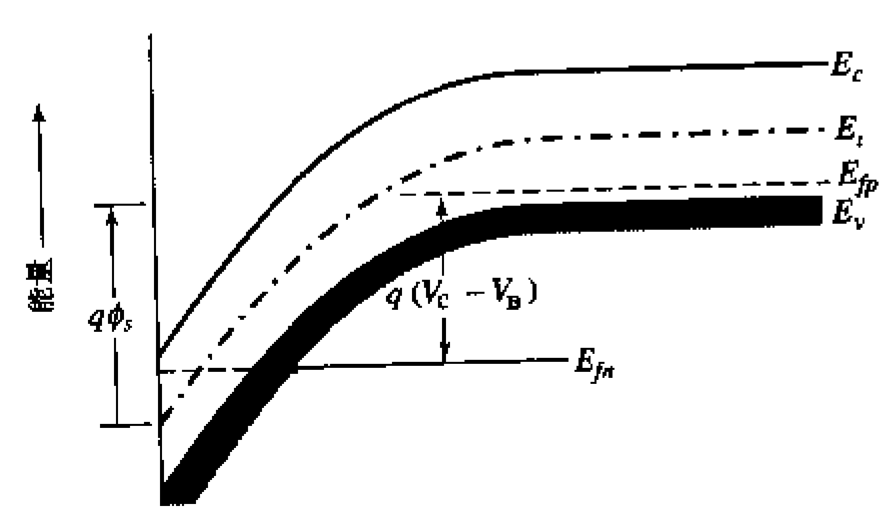


图 8.10　反型层和 p 型衬底之间加反偏电压($V_C - V_B$)时的能带图

表面反型层与衬底之间的反偏电压增加了耗尽层中的电荷 Q_d。因为由 $V_G - V_B$感应的负

电荷分布在耗尽层和反型层中,所以在栅极电压一定时,耗尽层电荷的增加意味着反型层电荷将减少。也就是说,当在表面反型层-衬底 pn 结上加反偏电压时,必须加更高的栅压才能在反型层中感应出相同数量的电子。外加反偏电压后使表面耗尽所需的栅压增大,使 x_{dmax} 增大。对表面反型层加反向偏置会使反型层的形成不如零偏时容易,因为该反向偏置的作用是排斥形成反型层的电子,只有当表面势 ϕ_s 达到 $|\phi_p|+V_C-V_B$ 时,而不是零偏时的 $|\phi_p|$ 时,才能形成反型层。因此,强反型条件变为 $2|\phi_p|+V_C-V_B$,而不是 $2|\phi_p|$。相应的最大耗尽层宽度 x_{dmax} 和耗尽层电荷 Q_d(单位面积)就变为

$$x_{dmax}=\sqrt{\frac{2\epsilon_s(2|\phi_p|+V_C-V_B)}{qN_a}} \tag{8.3.8}$$

和

$$Q_d=-\sqrt{2\epsilon_s qN_a(2|\phi_p|+V_C-V_B)} \tag{8.3.9}$$

前面关于 MOS 系统电荷的讨论表明,平带条件($V_G-V_B=V_{FB}$)对应于硅衬底的电中性状态。因此(V_G-V_B) $-$ V_{FB}是 MOS 电容的有效充电电压。本书中,MOS 系统的平带电压 V_{FB}类似于 pn 结的内建电势 ϕ_i;也就是说,两者都是对应于零偏压时的电荷存储状态,为了用外加电压表示电荷,我们进行如下分析:充电电压[(V_G-V_B) $-$ V_{FB}]是氧化层上的压降 V_{ox} 和硅中的压降($\phi_s-\phi_p$)之和,如图 8.6(d)所示。

$$V_G-V_B-V_{FB}=V_{ox}+\phi_s-\phi_p \tag{8.3.10}$$

在不考虑氧化层电荷时二氧化硅层中的电场保持为常数。用外加电压和氧化层厚度表示电场 $\mathscr{E}_{ox}$

$$\mathscr{E}_{ox}=V_{ox}/x_{ox}=[(V_G-V_B-V_{FB})-(\phi_s-\phi_p)]/x_{ox} \tag{8.3.11}$$

硅体内靠近氧化层处(不计入硅中电荷),电位移矢量的垂直分量 D 保持为常数,那么电场 $\mathscr{E}_{s0}$ 为

$$\mathscr{E}_{s0}=\frac{\epsilon_{ox}\mathscr{E}_{ox}}{\epsilon_s} \tag{8.3.12}$$

将等式(8.3.12)代入等式(8.3.11)中,利用氧化层电容(单位面积)的定义 $C_{ox}=\epsilon_{ox}/x_{ox}$,可得

$$\epsilon_s\mathscr{E}_{s0}=C_{ox}[(V_G-V_B-V_{FB})-(\phi_s-\phi_p)] \tag{8.3.13}$$

Gauss 定理表明包含在一定体积中的电荷等于离开该体积的电场与介电常数的乘积。对 Si-SiO_2界面到硅体内的中性区运用 Gauss 定理,可以写出

$$-\epsilon_s\mathscr{E}_{s0}=Q_s=Q_n+Q_d \tag{8.3.14}$$

式中半导体内感应的总电荷 Q_s等于可动电子电荷 Q_n与耗尽区电荷 Q_d(均为单位面积)之和。图 8.6(b)表示出了这些电荷。将等式(8.3.14)代入等式(8.3.13),得到可动电荷 Q_n的表达式为

$$Q_n=-C_{ox}[(V_G-V_{FB}-V_B)-(\phi_s-\phi_p)]-Q_d \tag{8.3.15}$$

将 ϕ_s和 Q_d的值代入等式(8.3.15),得到可动电荷 Q_n与外加电压的关系式。在沟道与硅衬底

间加反偏电压，且表面处于强反型时，$\phi_s = |\phi_p| + V_C - V_B$，此时等式(8.3.15)变为

$$Q_n = -C_{ox}(V_G - V_{FB} - V_C - 2|\phi_p|) + \sqrt{2\epsilon_s q N_a(2|\phi_p| + V_C - V_B)} \quad (8.3.16)$$

当沟道与衬底之间无偏置电压时，等式(8.3.16)简化为

$$Q_n = -C_{ox}(V_G - V_{FB} - V_B - 2|\phi_p|) + \sqrt{4\epsilon_s q N_a|\phi_p|} \quad (8.3.17)$$

值得注意的是，等式(8.3.16)和等式(8.3.17)等号右边第一项为负值而第二项为正值。然而，后者在数值上小于前者(因为$|Q_s| > |Q_d|$)，所以差值Q_n为负，这与p型衬底上的反型层电荷相对应。

由这些等式可以直接求出在半导体表面形成导电沟道所需的栅极电压，这个电压被称做阈值电压V_T，定义为在$Q_n = 0$时对应的栅极电压。由等式(8.3.16)，可写出V_T的表达式

$$V_T = V_{FB} + V_C + 2|\phi_p| + \frac{1}{C_{ox}}\sqrt{2\epsilon_s q N_a(2|\phi_p| + V_C - V_B)} \quad (8.3.18)$$

依次分析等式(8.3.18)中的各项，可以发现它们各自对应不同的作用。首先，V_T中的V_{FB}项是使衬底处于电中性(在图8.2所示系统中，V_{FB}为负，倾向于减小V_T)；其次，沟道电压V_C的增加导致需要更高的栅极电压才能使硅表面达到反型；再次，$2|\phi_p|$项指强反型时对应的表面电势；最后的平方根项是建立耗尽层电荷所需要的栅极电压部分。平方根项与氧化层电容成反比，并随$V_C - V_B$的增加而增加，它反映了$V_C - V_B$导致的半导体中电荷Q_s在反型层电荷(Q_n)与耗尽层电荷(Q_d)之间的再分布。

强反型时，反形层中的电荷可用栅极电压与阈值电压之差来表示，根据等式(8.3.16)和等式(8.3.18)，有

$$Q_n = -C_{ox}(V_G - V_T) \quad (8.3.19)$$

V_G非常接近于V_T时，等式(8.3.19)必须谨慎使用，因为它是在强反型(当表面势$\phi_s = -\phi_p + V_C - V_B$时)之前表面没有电子的假设下求出的。当然，对MOS系统精确求解Poisson方程时[1,2]可以不考虑这一假设。

MOS系统的常用公式见本章末的表8.3。表中同时给出了n型和p型衬底的结论。

8.4 MOS 系统的电容

类似于对Schottky势垒和pn结的研究，在输出电极间测量小信号电容，并对其进行分析，有助于研究MOS系统的电学特性。这个方法一直是研究$Si\text{-}SiO_2$系统及加工工艺的重要手段。MOS电容-电压(MOS *C-V*)的测量和分析技术是工艺开发以及生产中工艺诊断和监控的标准方法。本节将讨论MOS *C-V*特性的基本原理。这里我们考虑理想的MOS系统，假设氧化层中没有电荷，$Si\text{-}SiO_2$界面没有陷阱。

8.4.1 理想 MOS 系统的 *C-V* 特性

两端器件的小信号电容(单位面积)定义为两端电荷对所加电压的求导

$$C = \frac{dQ}{dV} \tag{8.4.1}$$

由等式(8.4.1)可看出,如果已知 Q 关于 V 的函数,可以很容易求出 C.正如前面讨论的 Schottky 二极管和 pn 结。此时,金属中的自由电子和硅中的多数载流子的介电驰豫时间均很短(一般在皮秒量级或更短),所以基本上可以即时响应外加电压的变化。然而,对于 MOS 结构,情况要复杂一些,电荷不能足够快地响应电压的变化,并且相关的电容会取决于电压变化的频率。后面将要讨论的各种情况总结在表 8.4 中。

表 8.4　MOS 系统中的电荷状态

偏置条件	电荷平衡状况	电荷分布	等效电路	说　明
低频(LF)	直流平衡 交流平衡	$x_d = x_{dmax}$, x, SiO_2, δQ_s	C_{ox}	$\delta Q_s = \delta Q_n$ 交流信号只与表面电荷耦合
高频(HF)	直流平衡 交流非平衡	$x_d = x_{dmax}$, SiO_2, δQ_B	C_{ox}　ϵ/x_{dmax}	$\delta Q_s = \delta Q_d$ 交流信号只与体电荷耦合
深耗尽(DD)	直流非平衡 交流非平衡	$x_d > x_{dmax}$, SiO_2, δQ_B	C_{ox}　ϵ/x_d	交流和直流信号都与 Q_d 耦合 $x_d > x_{dmax}$

首先考虑在稳定的直流偏置 V_G 下,硅表面处于积累状态的 MOS 系统。对于图 8.2 中所示的 p 型硅,对应的外加电压为负值,并且其电荷分布如图 8.4 所示。栅压所产生的静电力将硅表面的过剩空穴拉向氧化层。当在直流偏置 V_G 上叠加一个小的交流电压 v_G 时,它将分别在金属栅和硅表面引起存储电荷的微小变化,即 $-dQ$ 和 $+dQ$。假如系统中连入测量小信号电容的仪器,则测量出的结果几乎就是氧化层本身的电容,因为相对于氧化层的厚度,硅表面被调制电荷在空间上的扩展非常小。表面的积累程度越高,积累层越薄。因此,金属栅上的电荷 $-dQ$ 和硅中的电荷 $+dQ$ 间的距离接近于 x_{ox},因此其电容也接近于氧化层电容。这样,在积累状态下单位面积的电容 C 近似为

$$C_{ox} = \frac{\epsilon_{ox}}{x_{ox}} \tag{8.4.2}$$

式中 x_{ox} 是氧化层的厚度。当栅压接近于平带电压时,表面积累层消失,并且电容随着表面 Debye 长度的增加而减小。在这一偏置范围更精确的电容分析需要解 Poisson 方程[1],并要考虑表面附近对总空间电荷有贡献的自由电子、自由空穴以及掺杂原子。这里我们不仔细分析,只是定义 $V_G = V_{FB}$ 时测量的电容为 C_{FB}。根据参考文献[3],C_{FB} 可以表示成

$$C_{FB} = \frac{1}{1/C_{ox} + L_D/\epsilon_s} \tag{8.4.3}$$

式中 L_D 为非本征 Debye 长度，定义见等式(4.2.14)。

当栅极电压变得比平带电压更正的时候，空穴受排斥而远离硅表面，系统进入耗尽状态。在这一条件下，直接的分析表明，总电容 C 等于氧化层电容和表面耗尽层电容 C_s 的串联(习题8.3)。

$$C = \frac{1}{1/C_{ox} + 1/C_s} = \frac{1}{x_{ox}/\epsilon_{ox} + x_d/\epsilon_s} \tag{8.4.4}$$

式中 x_d 是表面耗尽层宽度，它取决于栅压、掺杂浓度以及氧化层性质。我们可以凭直觉得出等式(8.4.4)，而不需要考虑栅和硅中电荷 $\pm dQ$ 分布的详细推导。从物理上讲，这些电荷位于金属电极和硅中耗尽层的中性区边缘，类似平行板电容器的两个极板，被厚度为 x_{ox} 的氧化层和宽度为 x_d 的硅中耗尽层所隔开。总电容 C 为氧化层电容 C_{ox} 和耗尽层电容 C_s 的串联。由等式(8.4.4)可以看出，系统的电容随着耗尽区宽度的增加而减小。

当栅偏压增加使表面强反型时，描述 MOS 电容时就必须考虑其他的电荷。可以认为可动电荷位于硅中离 $Si\text{-}SiO_2$ 界面非常近的区域；这一反型层电荷对外加电压变化的响应能力决定了该偏置范围的电容特性。MOS 表面的反型层电荷是由少数载流子的产生形成的。因此，反型层中的电荷数量的变化只能与表面耗尽层中载流子的产生一样快。这一限制条件使得测量得到的小信号电容是外加交流信号频率 ω 的函数。

最简单的情况是直流栅偏置 V_G 和小信号电压 v_G 的变化均非常缓慢，此时，硅总是能达到平衡状态，并且反型层中的电荷数量完全跟得上外加信号电压的变化。在强反型状态下，栅和硅中的电荷仅仅被栅氧化层隔开，并且 MOS 系统的电容接近于 C_{ox}。在这些条件下，测得的电容与栅压的关系在图 8.11 中标为"LF"(低频)。积累区的电容值为 C_{ox}，并且随着表面耗尽电容值逐渐减小，表面强反型时又增加至 C_{ox}。

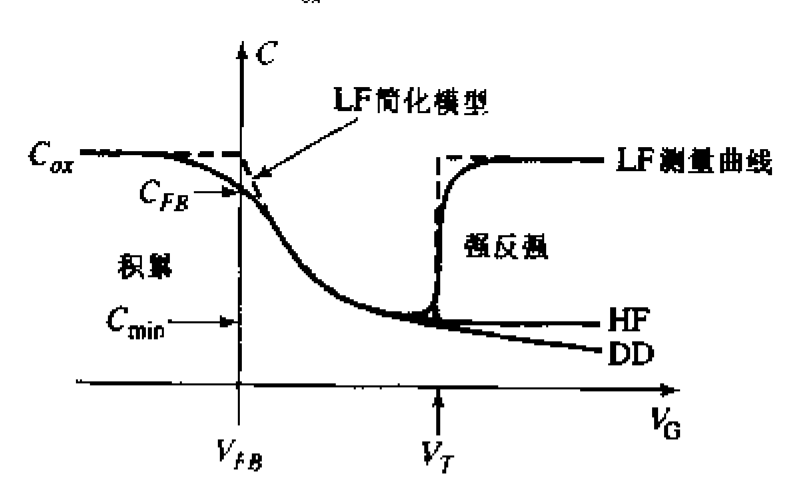


图 8.11　在不同的直流偏置和交流小信号条件下，理想 MOS 系统的 C-V 曲线。虚线表示对应于简化模型的低频(LF) C-V 曲线

习题 8.6 的结果表明，MOS 系统偏置在强反型状态时，形成表面反型层所需的特征时间 τ_{inv} 为秒量级或更长，并取决于表面少数载流子的寿命 τ_0($\tau_{inv} \sim 2N_a\tau_0/n_i$)。因此，小信号电压必须变化得非常缓慢才能够观察到低频 C-V 曲线。在本节后半部分讨论实际 C-V 测量技术时将考虑这一条件。

当直流偏置变化很慢而交流测量信号变化迅速时，反型层的变化能跟上直流偏置的变化，但是跟不上交流信号的变化。此时，电荷改变量 $-dQ$ 对应于远离表面的耗尽层边界处电荷的变化。电容则为氧化层电容和耗尽区电容的串联，类似耗尽偏置下的情形。由于当系统达到

强反型时耗尽区会达到最大宽度 x_{dmax}(等式(8.3.6)),所以此时测得的电容接近于最小值 C_{min},这个值等于氧化层电容与一个 x_{dmax} 宽度相对应的耗尽层电容的串联

$$C_{min} = \frac{1}{x_{ox}/\epsilon_{ox} + x_{dmax}/\epsilon_s} \tag{8.4.5}$$

当栅压进一步增加时,电容 C_{min} 将保持由等式(8.4.5)所给出的值。图 8.11 中也给出了高频(HF)C-V 曲线。

我们可以用 8.3 节中描述的简化模型来说明 C-V 曲线的特性。图 8.11 中的虚线给出了由此简化模型推导出的低频 C-V 曲线。简化模型假设积累、耗尽和强反型状态之间的转换是突变的,并且强反型层和积累层都无限薄。因此,这一模型高估了从积累到耗尽和从耗尽到强反型过渡时的电容。根据简化模型得出的低频 C-V 曲线给出的反型层电荷密度 Q_n 为

$$Q_n = C_{ox}(V_G - V_T) \tag{8.4.6}$$

等式(8.4.6)通常用来计算 MOSFET 导电沟道中的反型层电荷。这个表达式高估了 Q_n,因此也过高估计了 MOSFET 的输出电流。在按比例缩小的 MOS 工艺中,在低的偏置电压(即小的 $V_G - V_T$)下,这种过高估计电流的影响就更明显。第 9 章将继续考虑等式(8.4.6)。值得注意的是,MOSFET 的栅电荷和反型层电荷之间的电容起到了栅压与输运电流的反型层电荷之间的"耦合"作用;因为这一耦合作用的增强将提高晶体管的性能,所以这个电容是必要的。

图 8.11 中的"DD"曲线对应的电容被称为深耗尽电容。当栅偏置 V_G 和小信号电压 v_G 均以很高的速率变化时,表面耗尽区的产生-复合过程跟不上外加偏置的变化。由于不能形成反型层,导致耗尽层宽度大于 x_{dmax},这就是深耗尽的含意。此时的电容仍由等式(8.4.4)给出,但是,深耗尽时的 x_d 可以大于 x_{dmax},并且电容不会达到最小值。

8.4.2 C-V 测量中的实际因素

C-V 测量基础 C-V 测量系统是 MOS 工艺中研发和生产的标准工具。如图 8.12 所示,在典型测量系统中,加在 MOS 二极管上的电压由直流电压 V_G,叠加上由频率为 ω 的交流正弦小信号电压 v_G 所组成。为了在不同的偏置下进行测量,直流电压在一定范围内被设置成阶梯状增降的形式。通过二极管的电流和两端的电压通常由安培表和电压表来测量。从测量得到的阻抗 $Z = v/i$ 或者电导 $G = i/v$ 均可求出二极管电容 C。MOS 二极管可以等效为电容和(寄生)电阻 R_p 的并联或者串联的形式。在串联形式的电路中

$$Z = v/i = R_p + 1/j\omega C \tag{8.4.7}$$

在并联形式的电路中

$$G = i/v = 1/R_p + j\omega C \tag{8.4.8}$$

这样就确定了电容 C。

串联和并联模型显然将得到不同的电容和电阻值。通常情况下并联模型更有用,除非 MOS 二极管中的串联电阻比电容的阻抗值大很多。一般更高频率的交流电压能产生更大的电流,因此测量时有更高的信噪比和更高的精确度。但是,交流测量信号在通过电缆时存在延迟和相位变化的问题;这些问题随着频率的提高而趋于严重。假如不适当地考虑这些影响,由电缆带来的寄生效应可能导致对实验结果的错误解释。某些商用电容计有内部

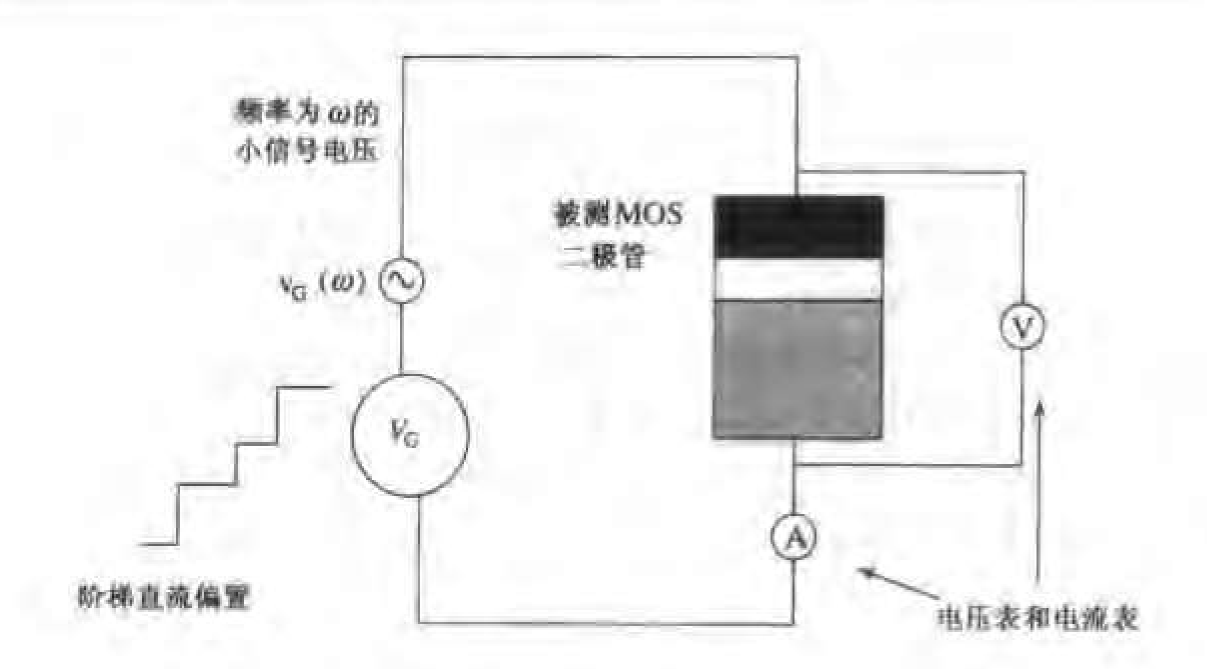


图8.12　MOS C-V测量系统。电压表和安培表测量加在二极管上的电压和流过二极管的电流

的校正和补偿措施，但是这些通常仅对特定的测量结构（例如，双线或四线）和特定的电缆长度有效。测量仪器的说明书中通常包括详细的测量方案以力求减小由于信号延迟导致的错误。

8.4.3　准静态（低频）C-V测量

低频C-V测量可以提取出下一节将要提到的界面陷阱密度的信息，这对工艺诊断和进程监控十分有用。前面已经描述了典型的低频和高频C-V曲线特征。我们注意到形成反型层的特征时间为秒量级或者更长，这使得采用常规的交流技术进行低频C-V测量不再实用。尽管通过光照可以增加样品中的电子-空穴对的产生率，从而得到类似于低频时的特性曲线。但是光照并不能够减少电子进出陷阱的时间。准静态C-V测量提供了更有用的低频C-V特性。图8.13给出了准静态C-V测量系统。准静态测量使用线性斜坡电压和静电计来测量位移电流，代替了阶梯状直流电压和叠加的交流测量信号。静电计是灵敏度（分辨率在fA量级）和速度很高的电表。当外加电压以R_R V/s的速率增加时，电路中的位移电

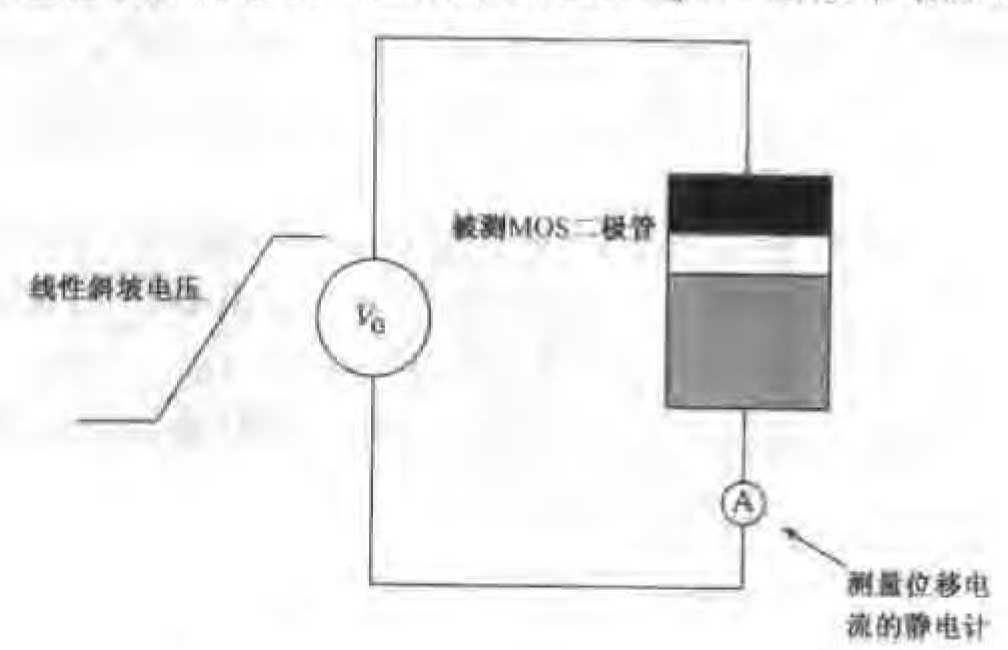


图8.13　测量低频C-V特性的准静态测量系统

流 I 可用 C 和 R_M 表示为

$$I = dQ/dt = C\,dV/dt = CR_m \tag{8.4.9}$$

对于 10pF 的电容,电压变化速率 $R_M = 0.1\text{V/s}$(适于测量低频 C-V 特性)时可产生 1pA 的电流,这个电流可以被精确地测量出来。

形成反型层所需要的时间较长,会导致测量低频[5] C-V 时出现问题,因为即使栅压变化非常缓慢(接近于直流)时,反型层也不能达到平衡。由这个过慢的产生过程引发的问题可以通过以下方法来避免:先在光照下产生载流子使电容偏置进入强反型,然后再去掉光照,测量栅压由强反型突变到积累型时的 C-V 特性。

8.5 非理想 MOS 系统

8.5.1 氧化层及界面态电荷

前面的讨论没有考虑氧化物-硅系统中的一个重要问题,即氧化层电荷和氧化物-硅界面态电荷的影响。但实际系统中这些电荷的存在是不可避免的。为了分析氧化层电荷的重要性,我们先估算 MOS 系统中的电荷密度的数量级。例如,当栅压略高于阈值电压,系统进入强反型区时,表面电子浓度 Q_n/q 与耗尽层的掺杂浓度(单位面积)有相同的数量级。对于均匀分布的杂质原子,当 N_a 为 10^{15}cm^{-3} 时,杂质原子的面密度为 $N_a^{2/3} \approx 10^{10}\text{cm}^{-2}$。与硅原子的面密度 $(5 \times 10^{22}\text{cm}^{-3})^{2/3} = 1.35 \times 10^{15}\text{cm}^{-2}$ 相比较,可以看出虽然表面电荷密度只是硅原子密度的 10^{-5},但却能够使得 MOS 系统偏离其理想特性。幸运的是,如果工艺控制得当,热生长的无定型二氧化硅和单晶硅界面中的电荷密度仅为 10^{10}cm^{-2} 量级或更低。界面电荷密度低是硅-二氧化硅系统中的一个显著特性。下面先分析氧化层电荷对 MOS 系统的影响,然后考虑它们具体的起源。

理论分析 考虑电荷密度 Q_{ox} 位于氧化层中 $x = x_1$ 平面处,如图 8.14(a)所示。x_1 处的电荷在硅和金属栅上感应等量的异种电荷。x_1 越接近 Si-SiO_2 的界面 x_{ox},则在硅中感应的电荷所占的比重就越大。由于这些感应电荷改变了热平衡状态下硅中存储的电荷,于是平带电压也就偏离了在理想 MOS 分析时所得到的值(等式(8.1.1))。图 8.14(a)对应的零偏压情况下电场和电势分布如图 8.14(c)所示;图8.14(b)对应的平带情况下的电场和电势分布如图 8.14(d)所示。

由氧化层电荷引起的平带电压的改变量可以用 Gauss 定理求出。当栅电极上感应了与氧化层电荷 Q_{ox} 等量的镜像电荷而硅中不产生感应电荷时,可求出由 Q_{ox} 引起的相应的栅电压,这就是平带电压的改变量。这一情形见图 8.14(b)。由该图和 Gauss 定理,可以看出在平带情况下(见图 8.14(d)),金属($x = 0$ 处)和 Q_{ox}(x_1 处)之间的电场保持常数,而 x_1 和硅表面($x = x_{ox}$ 处)之间的电场为零。栅和 x_1 之间的电场 $\mathscr{E}_{ox}$ 为

$$\mathscr{E}_{ox} = -\frac{Q_{ox}}{\epsilon_{ox}} \qquad 0 < x < x_1 \tag{8.5.1}$$

---

5 译者注:原文误为"高频"。

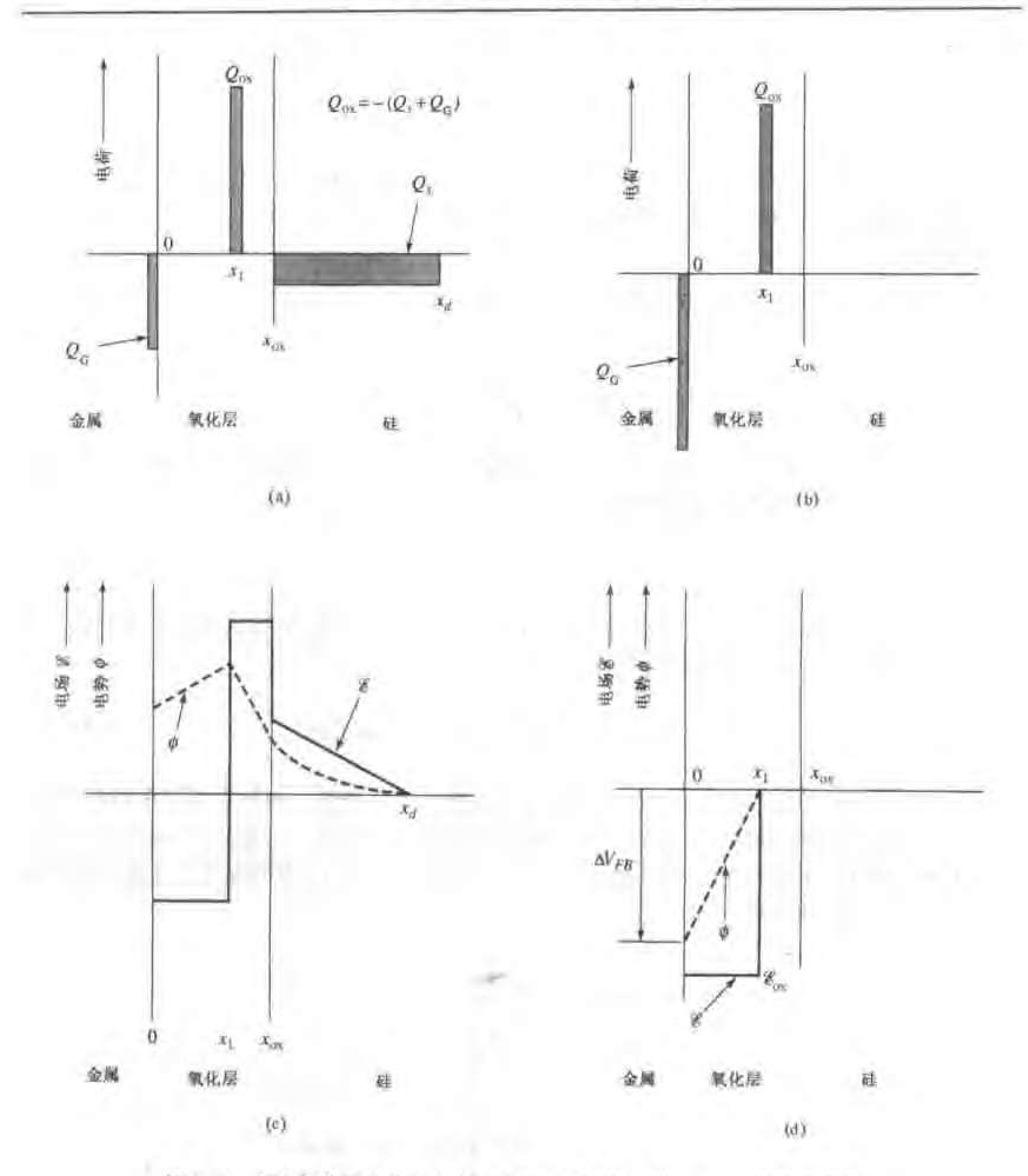


图 8.14　固定氧化层电荷密度 Q_{ox} 对 MOS 系统的影响。(a)零偏置时的电荷分布，$Q_{ox}=Q_s+Q_G$；(b)平带时的电荷分布，$Q_{ox}=Q_G$；(c)零偏置时电场(实线)和电势(虚线)分布；(d)平带时电场(实线)和电势(虚线)分布。(c)和(d)中取体硅为电势参考点

由 Q_{ox} 引起的栅电压等于 $\mathscr{E}_{ox}$ 在氧化层中积分的负值。由于它对平带电压有贡献，所以称之为 ΔV_{FB}(即相对于理想 MOS 情形，由氧化层面电荷引起的 V_{FB} 的改变)。ΔV_{FB} 的表达式为

$$\Delta V_{FB}=x_1\mathscr{E}_{ox}=-\frac{x_1 Q_{ox}}{\epsilon_{ox}} \tag{8.5.2}$$

等式(8.5.2)可以用等式(8.4.2)中单位面积氧化层电容 C_{ox} 重新写成如下形式

$$\Delta V_{FB} = -\frac{Q_{ox} x_1}{C_{ox} x_{ox}} \tag{8.5.3}$$

当 Q_{ox} 位于 Si-SiO$_2$ 界面($x_1 = x_{ox}$)时 ΔV_{FB} 取最大值,因为此时由 Q_{ox} 感应的电荷全部分布在硅中。相反,当 Q_{ox} 靠近栅电极($x_1 = 0$)时,感应电荷全部分布在金属中,氧化层电荷对 ΔV_{FB} 没有任何作用。

上述结论可以推广到任意分布的氧化物电荷 $\rho(x)$ 对平带电压的改变,此时平带电压的改变量为

$$\Delta V_{FB} = -\frac{1}{C_{ox}} \int_0^{x} \frac{x}{x_{ox}} \rho(x)\,dx \tag{8.5.4}$$

尽管等式(8.5.4)中也包含了表面电荷,我们通常将 Si-SiO$_2$ 界面的固定电荷与氧化层内部的电荷分开考虑。固定的界面电荷密度用 Q_f 来表示,它对平带电压的贡献为

$$\Delta V_{FB} = -\frac{Q_f}{C_{ox}} \tag{8.5.5}$$

将等式(8.1.1)、等式(8.5.4)以及等式(8.5.5)相结合,可得到 V_{FB} 的表达式。式中考虑了硅和栅电极功函数差以及氧化层固定电荷的作用

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{x_{ox}} \frac{x}{x_{ox}} \rho(x)\,dx \tag{8.5.6}$$

等式(8.5.6)表明氧化层电荷使平带电压偏离了理想状况下的值。如果氧化层电荷稳定,则平带电压的改变将使阈值电压 V_T(式(8.3.18))产生相应的变化。实验证实,阈值电压的改变使电容与栅压 C-V 曲线沿 V_G 轴平移,但不改变其形状。图 8.15 中的虚线为考虑氧化层电荷时典型的高频电容曲线。

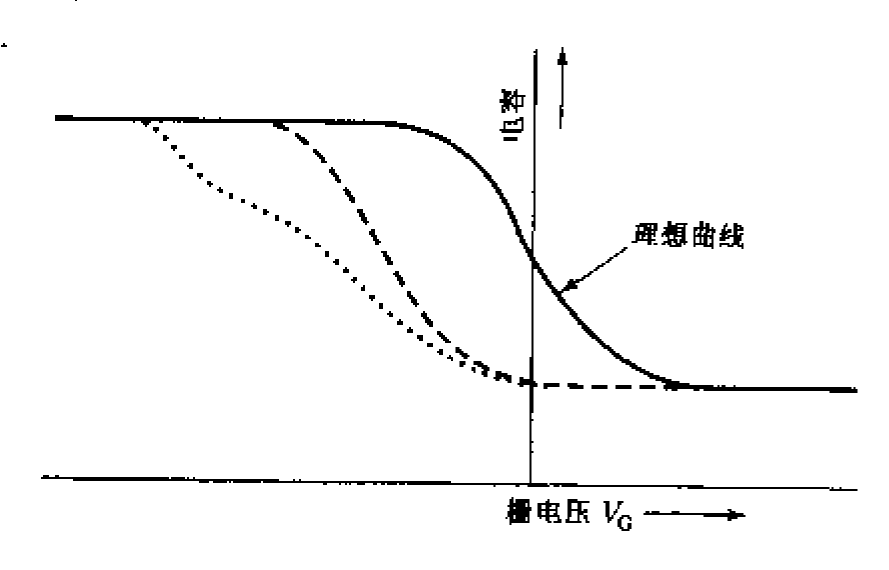


图 8.15 氧化层中的固定电荷引起 C-V 曲线沿 V_G 轴平移但不改变形状(虚线);电荷受栅压影响时将引起 C-V 曲线畸变(点线)

在某些情况下,氧化层以及 Si-SiO$_2$ 界面包含的电荷会受到所加电压的影响而变得不稳定,这时的阈值电压本身与栅电压有关,C-V 曲线将发生畸变,如图 8.15 中的点线。为了解释这一现象,并研究固定电荷的影响,下面将讨论氧化层电荷的起源。

8.5.2　氧化层电荷的起源

下面讨论 Si-SiO_2系统中的四种不同类型的电荷。图 8.16(a)画出了这四种类型的电荷，同时标注了普遍采用的名称以及符号[4]：Q_f为固定电荷；Q_{ot}为氧化层陷阱电荷；Q_{it}为界面陷阱电荷；Q_m为可动电荷。

前面已经讨论了固定界面电荷 Q_f(相应的固定电荷数 $N_f = Q_f/q$)。这类电荷为正，位于 Si-SiO_2间非常薄(< 1nm)的非化学剂量比的硅氧化物(记为 SiO_x)过渡层中，如图 8.16(a)所示。氧化层陷阱电荷 Q_{ot}可以为正也可以为负(通常为负)，位于遍布于整个氧化层的陷阱中。制备过程中引入的氧化层陷阱电荷非常少，通常可被忽略。除了在某些特殊的条件下(第 10 章中讨论)，这些电荷一般为固定电荷。我们先考虑图 8.14(a)中所示的第四种电荷，即位于界面陷阱 N_{it}上的界面陷阱电荷 Q_{it}，然后再讨论可动电荷。

界面陷阱 N_{it}(陷阱数 cm^{-2})位于氧化物-硅界面，类似第 5 章讨论的陷阱，其能级位于禁带中(图 8.16(b))，以态密度 D_{it}($cm^{-2}eV^{-1}$)分布在禁带中。3.5 节在讨论金属-半导体接触时，从我们对硅表面引入的允许能级的各种来源的讨论中可知，即使清洁的表面也会引入不同于晶体体内的允许能级。在芯片制造过程中不可避免地会引入杂质和晶体缺陷，它们是引入的允许能级的来源。这些能级上的电子和与之相关的离子都对界面电荷有贡献。

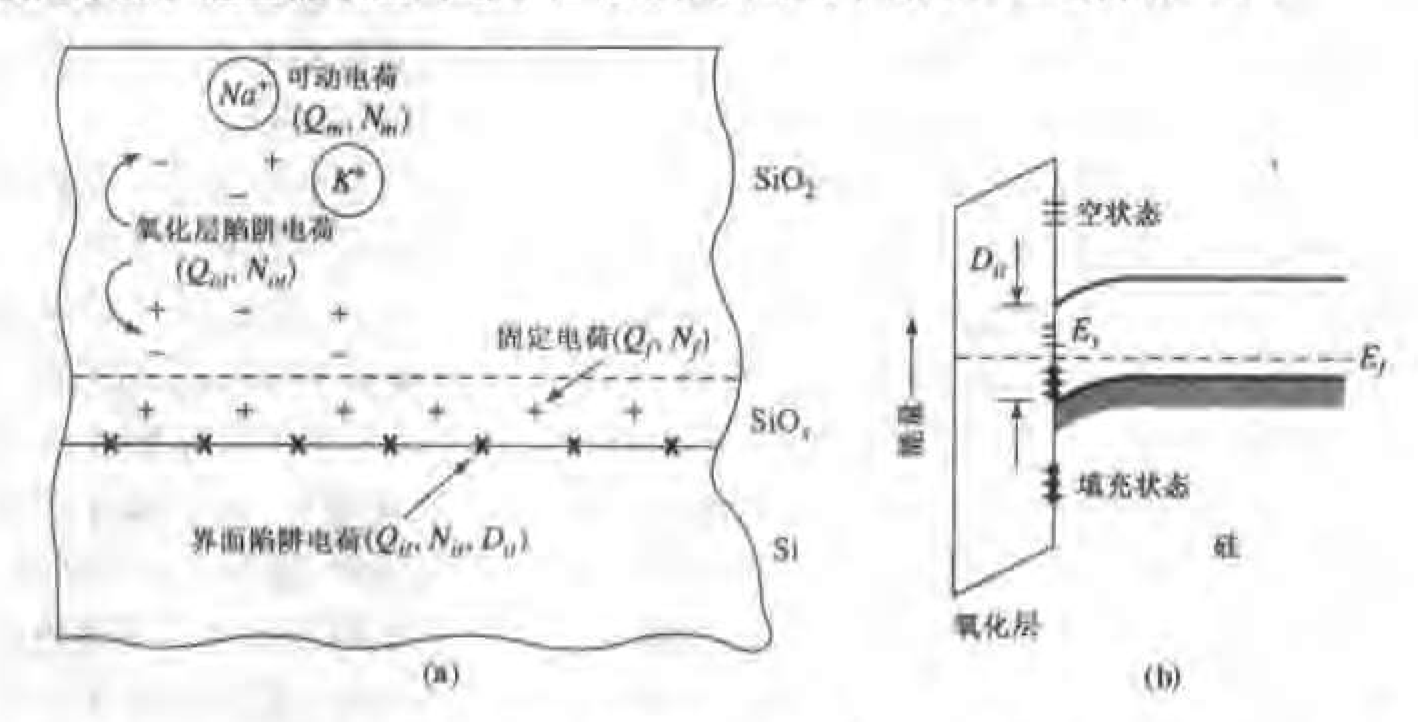


图 8.16　(a)MOS 系统氧化层中的四类电荷。标准符号是，电荷密度用 Q(C/cm^{-2})表示，态密度用 N(状态数 cm^{-2})或 D(状态数 $cm^{-2}eV^{-1}$)表示[4]；(b)氧化物-硅界面处的能级。界面陷阱能级用禁带中的态密度 D_{it}(状态数 $cm^{-2}eV^{-1}$)表示

为了将图 8.15(点线)所示的畸变 C-V 曲线与这些陷阱的作用联系起来，现考虑图 8.16(b)中氧化物-硅界面处位于 E_t能级的界面陷阱。如果栅压使 Fermi 能级在表面与 E_t相交，则这些陷阱的荷电状态必然会发生改变。界面陷阱电荷密度的改变在等式(8.5.6)中引入了一个与电压有关的项 Q_{it}/C_{ox}，使得平带电压和阈值电压(等式(8.3.18))均随 V_G而变化，并且导致 C-V_G曲线的畸变。为了使器件性能可靠，如果界面陷阱密度接近于典型的反型层电荷密度($10^{10}cm^{-2}$量级或更高)，一般是不能接受的。多年来生产中不能降低较高的陷阱电荷密度，是造成 MOS 器件的实际应用落后于理论 30 年的主要原因。但是，现代 MOS 工艺可以将这些陷阱电荷密度减小到可接受的程度。典型的方法是在氢气或形成气体(氢气和氮气的

混合物)中对氧化后的硅片进行退火

界面处的固定电荷(总是存在)(Q_f)和氧化层中的固定电荷(较少)可以改变阈值电压。因此,对实际的电路工作,高的固定电荷密度会造成阈值电压过高,尤其是在外加电压减小时。界面处总是有一些固定电荷(Q_f)存在,这是由不饱和的硅-硅键引起的。硅晶体表面处的原子密度与晶向有关,因此Q_f也与硅片的晶向有关。与(100)硅片相比,(111)硅片在形成二氧化硅时需要断开更多的硅-硅键,因此Q_f一般会大一些。正是由于这一原因,几乎所有的商用MOS器件都制作在(100)硅片上。另外,固定电荷密度也与工艺中的高温处理过程有关,尤其是最后的工艺步骤。通过高温退火可使一些断裂的键重新成键(Q_f减少)。

图8.16(a)中所示的可动电荷Q_m通常是二氧化硅中吸收的碱金属离子(主要是钠和钾)。钠极其广泛地分布在许多金属和化学物质中,而且非常容易通过人为的接触进人氧化层。由于这些离子是带电的,根据等式(8.5.4),其位置和总浓度都会影响V_{FB}。碱金属离子在较低的外加电压下,就能够在氧化层中漂移。随着温度的升高,碱金属离子的迁移率迅速增大,所以高温下平带电压的不稳定的问题尤为突出。由于金属离子带正电,负栅压可以使这些离子移动到金属-氧化层界面,从而不对平带电压产生影响。但是,正栅压将使得这些离子移动到氧化物-硅界面,此时对V_{FB}的影响最大。总之,氧化层中的可动离子使MOS结构性能不稳定。若阈值电压的波动限制在0.05V(50mV)范围,则对于10nm厚的氧化层,可容忍的可动离子应少于$1\times10^{11}\text{cm}^{-2}$。可动离子污染是制约实际MOS系统发展的严重问题。通过控制制备工艺和引入杂质来固定碱金属离子可以避免这一问题。在热氧化过程中引入Cl及其化合物,尤其是HCl,可以制备稳定的MOS结构(尽管这些添加剂会带来其他问题)。

辐射也有可能在MOS系统中引入电荷,这些辐射源可以是制造器件过程中入射的高能电子,或者是空间环境中的高能粒子或光子,或者是由封装材料发射的高能粒子或光子。辐射可以改变氧化层电荷Q_{ot}和界面陷阱态密度D_{it}。例如,高能光子(即能量高于二氧化硅的带隙8eV或9eV的光子)可以在氧化层中产生电子-空穴对,光子能量大于硅带隙宽度的光照也可以在硅中产生电子-空穴对。但是,氧化层中产生的电子-空穴对不同于半导体中产生的载流子。由于氧化层非常薄,而且几乎没有自由载流子,所以复合几率很小,而大部分产生的电子被加在氧化层上的电场扫出。但氧化层中通常含有大量的空穴陷阱,很多空穴被固定,从而增加了氧化层中的正电荷。正如等式(8.5.4)指出的,陷阱对空穴的俘获改变了平带电压。

从图8.2中可以看出,能量比二氧化硅带隙低的光子可以将电子激发到氧化层中。如果光子的能量超过了能量势垒(铝一侧为3.15eV,硅一侧如从导带发射时为3.1eV),则这些电子可以从金属或半导体中被光致激发到二氧化硅。实际上,只有光子能量足以将电子从价带(电子数量较多)中激发时(要求光子能量高于4.2eV),硅中才出现光激发现象。如果氧化层因为存在被陷空穴而带电,则将电子从金属或者硅光致激发到氧化层时,部分光致激发电子将与被陷空穴复合,从而减少了氧化层中的正电荷。

由4.4节知,当硅中出现雪崩击穿时会产生高能电子,因此,雪崩击穿是使电子获得高于氧化层界面势垒的能量的另一种途径。这种机制受外加电压的控制,因而对器件应用尤其重要。第10章将更全面地考虑利用雪崩注入的器件。就目前的讨论,我们可以简单地认为,硅表面附近的雪崩击穿是除光致激发外提供进入氧化层的高能电子的另一种途径。

总之,上述各种类型的电荷Q_f、Q_{ot}、Q_{it}和Q_m都能影响平带电压(等式(8.5.6))和阈值电压(等式(8.3.18))。这些表达式中只明确考虑了固定电荷密度Q_f,其他电荷由于密度很低,

如要考虑它们的作用，可包含在 V_{FB} 表达式的对氧化层电荷的积分式中。

8.5.3 氧化层电荷的实验测定

上述讨论表明，MOS 系统中的氧化层电荷可以改变其电学特性。固定电荷可改变平带电压，而界面陷阱电荷会影响到硅表面从耗尽区向反型区的过渡，导致阈值电压改变。下一章将看到，带电的界面陷阱还会减小反型层载流子的迁移率，进而减小 MOS 晶体管的漏极电流。为确保 MOS 工艺能制造稳定可靠的器件，工艺工程师需要改进和开发工艺，以尽可能地减小氧化层中的电荷，MOS 制备过程中要不断监测氧化层中电荷的数量及类型。一种广泛使用的测量氧化层中可动电荷的技术称做"偏压-温度"法：在高温环境下(例如 125℃)对 MOS 结构施加电场来移动二氧化硅中的可动离子，改变它们对 V_{FB} 的影响；然后在室温下测量其 C-V 曲线。最后在高温下反方向加偏压，并在室温下再次测量 C-V 曲线。C-V 曲线的迟滞表明了氧化层中可动电荷的数量，并且曲线的具体形状提供了被陷电荷(见习题 8.14)的信息。最大迟滞为几十毫伏时对应的可动电荷密度通常是允许的。

如图 8.17 所示，定量测定界面陷阱密度一般有两种不同的方法。第一种方法是测量 MOS 电容的高频(通常为 100kHz 或者更高)C-V 曲线，并与理想 MOS 系统(没有氧化层电荷)的理论 C-V 曲线作比较。为了获得理论 C-V 曲线，必须知道精确的杂质浓度、杂质分布以及氧化层厚度。这些参量通常用物理、光学及电学相结合的方法，在器件制造过程中和制备完成后测量得到。为了说明这一方法，假定已经获得实验和理论 C-V 曲线并进行了适当的校准。在强反型开始之前的区域，考虑两曲线上电容值相同的两个点(图 8.17(a))。由于实验曲线是在高频下得到的，此时界面态跟不上交流信号。因此，这两点对应的耗尽层电容和耗尽层宽度(以及表面势)是相等的。但是对于测量得到的非理想电容，能量低于 Fermi 能级的界面陷阱均被电子填充，所以理论和实验之间产生了栅电压的偏差。这个电压差 $V_{G2}-V_{G1}$，可以用界面态密度 D_{it}(它与能量有关，因此与 ϕ_s 有关，单位为 $cm^{-2}eV^{-1}$)表示出来

$$V_{G2}-V_{G1}=\frac{q\int_{\phi_{s1}}^{\phi_{s2}}D_{it}(\phi_s)d\phi_s}{C_{ox}} \tag{8.5.7}$$

式中 ϕ_{s1} 和 ϕ_{s2} 是分别对应于 V_{G1} 和 V_{G2} 的表面势。表面势 ϕ_s 的值可以从理想 C-V 曲线中求出，

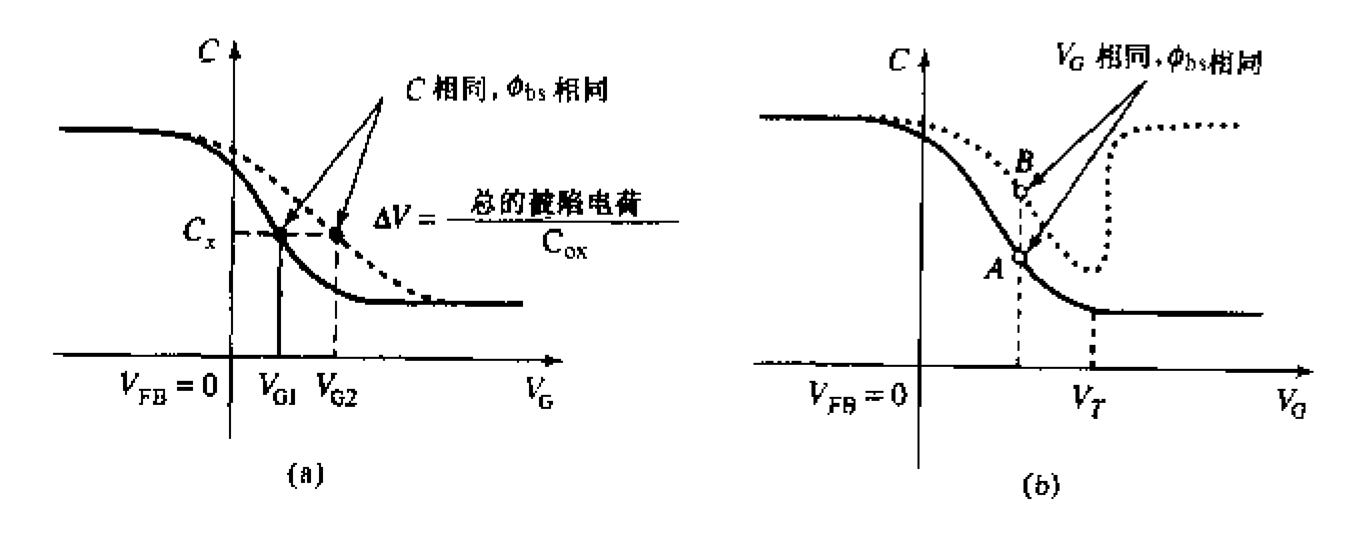


图 8.17　测量界面陷阱密度的两种方法：(a)理想和非理想高频 C-V 曲线的特性；(b)非理想低频和非理想高频 C-V 曲线的特性

利用这些电容值,可用数值方法求出 D_{it} 与 ϕ_s 的函数,从而得到带隙中界面陷阱的态密度分布。

第二种方法是将同一电容测量的高频 C-V 曲线与测量的低频 C-V 曲线相比较(图 8.17(b))。低频 C-V 曲线通常由前面提到的准静态方法获得。由于这两条曲线的直流偏置条件相同,所以相同的栅偏压对应于相同的表面势。由于高频 C-V 测量时界面陷阱跟不上信号的变化,所以等效电容可以简单地看做是氧化层电容 C_{ox} 和硅耗尽层电容 $C_{sHF}=\epsilon_s/x_d$ 的串联

$$C_{HF}=\frac{1}{1/C_{ox}+1/C_{sHF}}=\frac{1}{x_{ox}/\epsilon_{ox}+x_d/\epsilon_s} \tag{8.5.8}$$

低频下,衬底中的电容还包含界面陷阱充放电作用的影响

$$C_{LF}=\frac{1}{1/C_{ox}+1/C_{sLF}} \tag{8.5.9}$$

式中

$$C_{sLF}=\frac{dQ_s}{d\phi_s}=\frac{d(Q_{it}+Q_d)}{d\phi_s}=qD_{it}+\frac{\epsilon_s}{x_d} \tag{8.5.10}$$

利用等式(8.5.8)、等式(8.5.9)和等式(8.5.10),D_{it} 可由相应的表面势 ϕ_s 计算出。不同栅压下 ϕ_s 的值可用习题 8.5 中介绍的 Berglund 方法来确定[5]。实际的 MOS 系统,界面陷阱密度通常在带隙中成 U 形分布,靠近带边的值高于带中部的值,如图 8.18 所示。

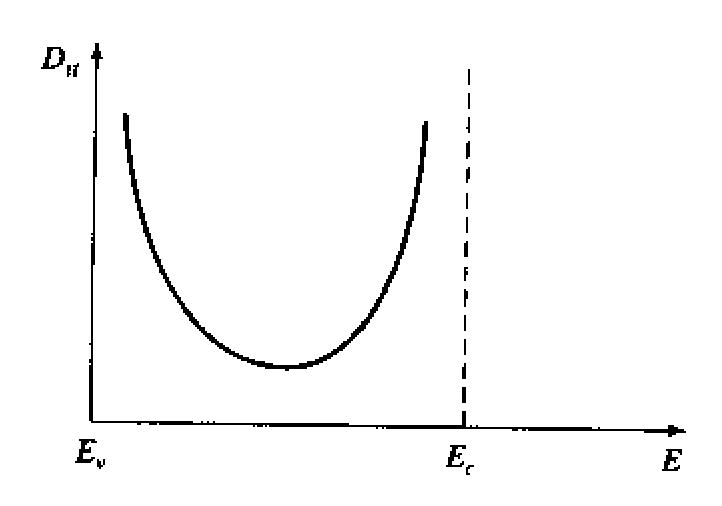


图 8.18 氧化硅-硅界面处硅禁带中界面陷阱态密度的 U 形分布

8.6 pn 结的表面效应†

考虑图 8.9 所示的结构,当被栅覆盖的氧化层附近存在 pn 结时,将出现一些影响器件工作的重要效应。我们在 8.3 节讨论图 8.9 时已经看到,对于反型的 MOS 系统,加在结上的偏置电压可以改变存储在沟道的电荷 Q_n 和存储在耗尽区的电荷 Q_d。这是 MOS 系统的结效应。集成电路设计者也迫切需要了解 MOS 系统对结的性质的影响。多数由平面工艺制造的 pn 结与氧化物-硅界面相交。因此,氧化物-硅系统的性质将对 pn 结器件的电路性能产生重大影响,例如双极晶体管。

图 8.19 是平面 n^+p 型集成电路二极管的横截面示意图。如果这个二极管位于另一个 pn 结之上,它将代表双极晶体管中的发射结。5.3 节指出 pn 结耗尽区中电子-空穴对的产生通

常是 pn 结反向漏电流的主要来源。类似地，在低的正向偏压下，空间电荷区中的复合电流是正向电流的主要分量。空间电荷区中的产生与复合不仅使电流电压关系偏离理想二极管定律，更严重的是，这些过程增加了不被集电极收集的基极电流分量，从而降低了晶体管的性能（见第6章）。发生在表面空间电荷区的这些特殊效应值得我们重视。

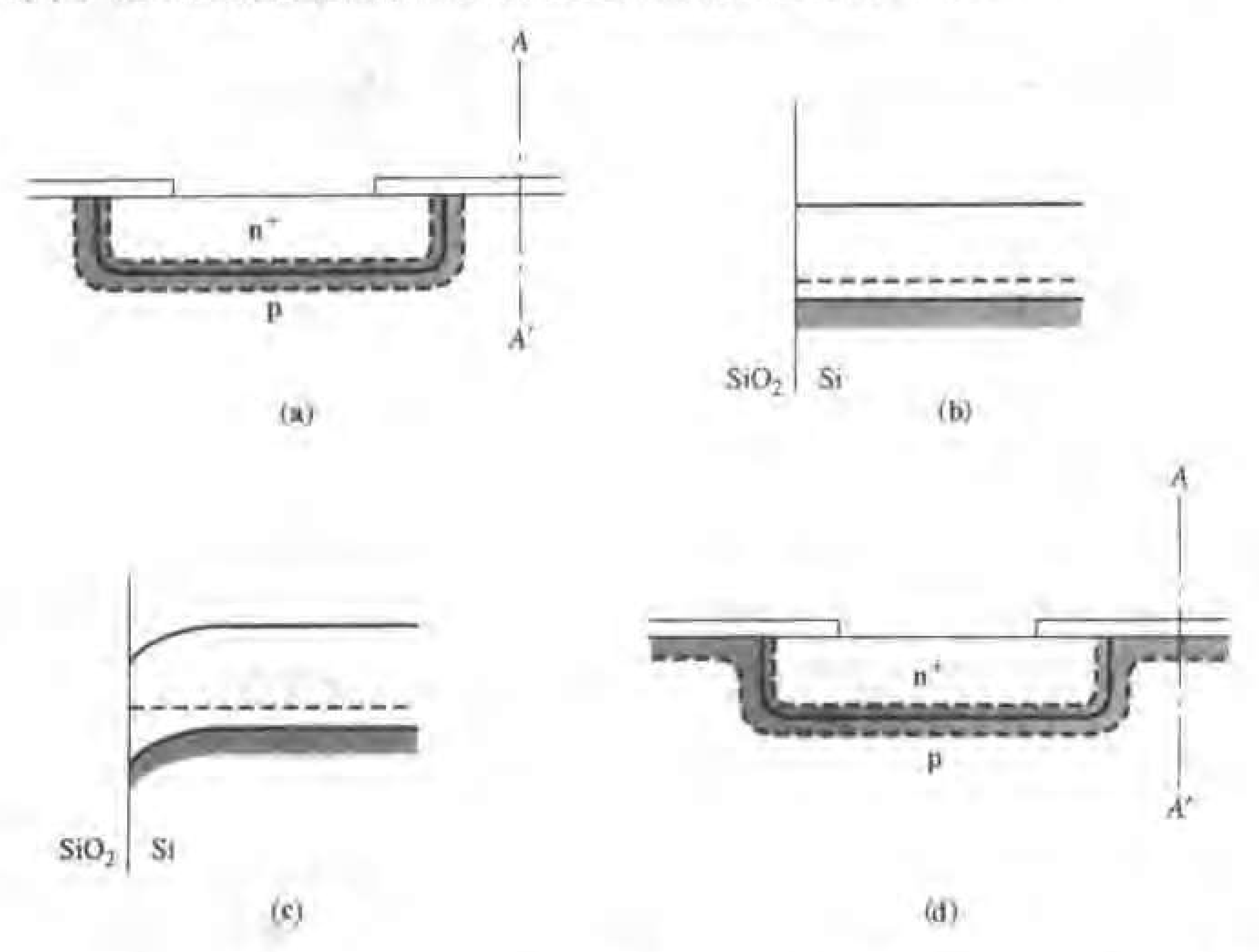


图 8.19　(a) 平面二极管；(b) 理想状态下沿 $A-A'$ 方向为平带；(c) 由于正的氧化层电荷，实际的氧化层 - p 型硅界面处于耗尽状态；(d) 表面耗尽区与结耗尽区相连

考虑图 8.19(a) 中通过二极管的 $A-A'$ 截面。理想 pn 结分析认为氧化物-硅界面不带电荷并且处于平带状况，如图 8.19(b) 中所示。但是，对 MOS 系统的分析表明平带条件并不对应于热平衡。为达到平带条件，需要在栅和衬底之间加一个等于 V_{FB} 的电压。假如在 $A-A'$ 截面处的氧化层上方没有电极，则此时对表面产生的影响最大的是氧化层电荷。氧化层电荷通常为正电荷，引起 p 型硅表面耗尽甚至反型（图 8.19(c)）。相反，氧化层中的正电荷趋向于使 n 型硅中表面出现多子积累（习题 8.13 讨论了在双极晶体管的制备中，这些效应所产生的影响）。我们再来考虑图 8.19 中所示的 p 型区，可以看出在 $A-A'$ 截面附近的表面耗尽区使总的结耗尽区增加，因为此时表面耗尽区与扩散结的耗尽区相连。假如表面处于反型，则沿氧化层表面的导电 n 型区将扩展，耗尽区也将增大。

耗尽区扩展带来的一个效应是反向偏置下空间电荷区体积增大，因而导致产生流的增大。更重要的是，氧化物-硅界面产生-复合中心的密度比体内的高。另外，这些复合中心的有效性取决于氧化物-硅表面的表面势，这与 SHR 理论（见 5.2 节）所揭示的相一致。

栅控二极管结构[†]

为了研究表面势对 pn 结的影响，我们引入栅控二极管的概念[6]。这一结构如图8.20(a) 所示，金属栅覆盖了表面处 pn 结两侧的 p 型区与 n 型区。为了研究栅控二极管的物理机制，

我们将体硅接地($V_B=0$),并在 n 型区加正电压 V_R(结为反偏)。

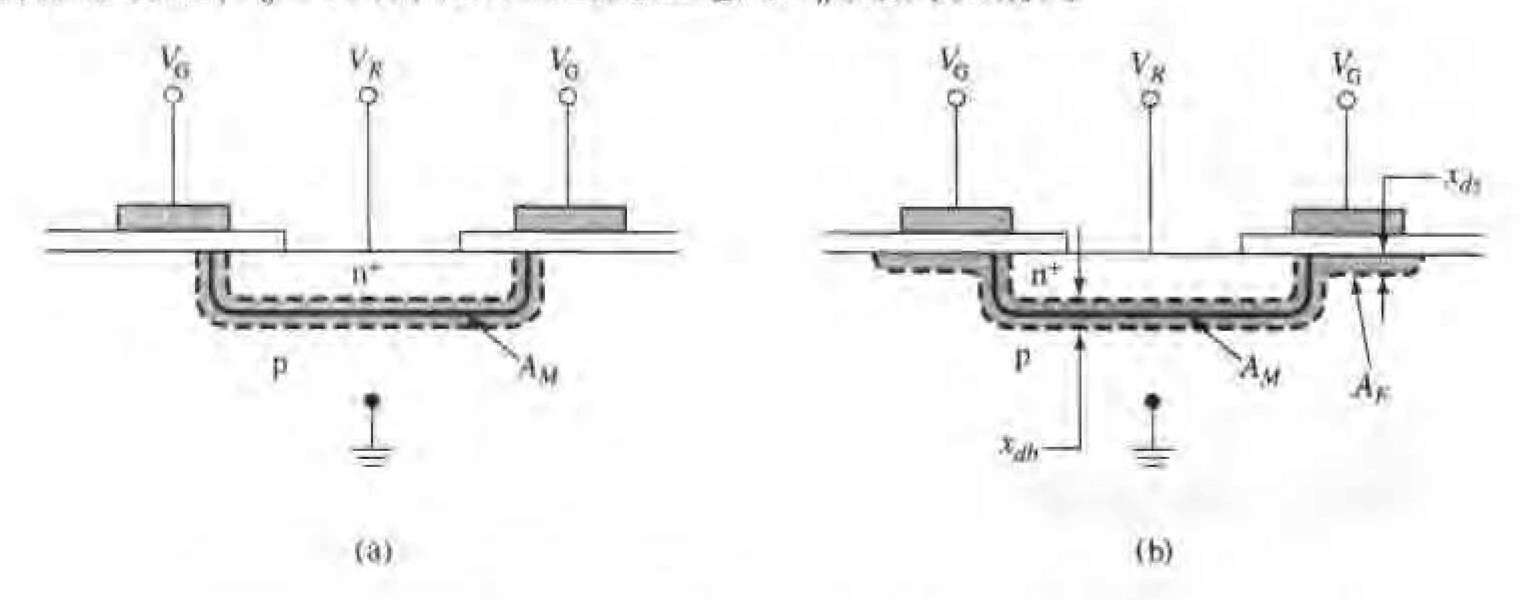


图 8.20 标出了耗尽区的栅控二极管结构(a)p 型区栅压等于平带电压 V_{FB},(b)栅压导致硅表面耗尽区宽度为 x_{ds}

当栅压 V_G 相对于平带状态为负时,p 型区表面为多子积累且表面耗尽区很小。n 型区由于高掺杂,其表面附近的耗尽区几乎不受栅压控制。因此,对于负 V_G,二极管的漏电流大致由冶金结区域的空穴和电子的产生率所决定。由等式(5.3.26),产生电流 I_M 为

$$I_M=\frac{qn_i}{2\tau_0}x_iA_M \tag{8.6.1}$$

式中 τ_0 为寿命,A_M 为冶金结面积,x_i 是 5.3 节描述的载流子产生的有效区。实际上,可以认为 x_i 与体内耗尽层宽度 x_{db} 相等,因此 x_i 与 x_{db} 以相同的方式随 V_R 变化(通常为$(V_R+\varphi_i)^{1/2}$)。在这种偏置条件下,电流几乎不受 V_G 的影响。

如果 V_G 相对于平带电压为正,则栅电极下的 p 型区表面就会出现耗尽区,如图 8.20(b)所示。此时,除了等式(8.6.1)中给出的 I_M 外,二极管反向电流中还包含另外两部分。首先是由栅引入的耗尽区中载流子的产生流。由等式(5.3.26)可求出这部分电流,我们称之为 I_F(由场致结产生)

$$I_F=\frac{qn_i}{2\tau_0}x_{ds}A_F \tag{8.6.2}$$

式中 x_{ds} 为表面耗尽区的宽度,是外加栅压 V_G 的函数;而 A_F 为表面耗尽区的面积,取决于栅电极对 p 型区的覆盖程度。p 型区耗尽时的第二个电流分量由表面的载流子产生形成,该电流称为 I_S。用 5.2 节介绍的表面复合速度 s 可以很好地表示这一电流,而表面复合速度正比于该处的产生-复合中心密度 N_{st}(等式(5.2.23))[⑥]。如果认为这些复合中心的能量接近于 E_i,则(当 p_s 和 $n_s \ll n_i$ 时利用式 5.2.20)可以求出 I_S 为 q 乘以产生率

$$I_S=\frac{qn_is_0A_F}{2} \tag{8.6.3}$$

式中 s_0 的定义见等式(5.2.24),$s_0=N_{st}v_{th}\sigma$,v_{th} 是热运动速度,σ 是产生-复合中心相关的俘获

⑥ N_{st} 是 8.5 节介绍的界面陷阱态密度 N_{it} 的特例。N_{st} 态的能量位于本征 Fermi 能级的附近,并与电子和空穴有几乎相等的交换率。

截面。s_0的值正比于表面产生-复合中心的数量，所以器件工艺及退火过程对它的影响极大。

如果进一步增加栅压直到 p 型硅表面反型，等式(8.6.2)中的 I_F 也将增加，并且在 $x_{d0}=x_{dmax}$时达到最大值。但是反型层一旦出现，表面电子浓度 n_s就远大于本征浓度 n_i，并且根据等式(5.2.23)，表面复合速度 s_0显著减小。当表面耗尽时，I_F的典型值(见习题8.16)小于 I_S。反型后的 I_S(表面产生电流)可以忽略，则反偏电流变为 I_M和 I_F的和。图 8.21 中绘出了栅控二极管反向漏电流的典型测量特性。当 pn 结低为正向偏置时，表面空间电荷区也会影响复合电流，但是这个影响一般很小。

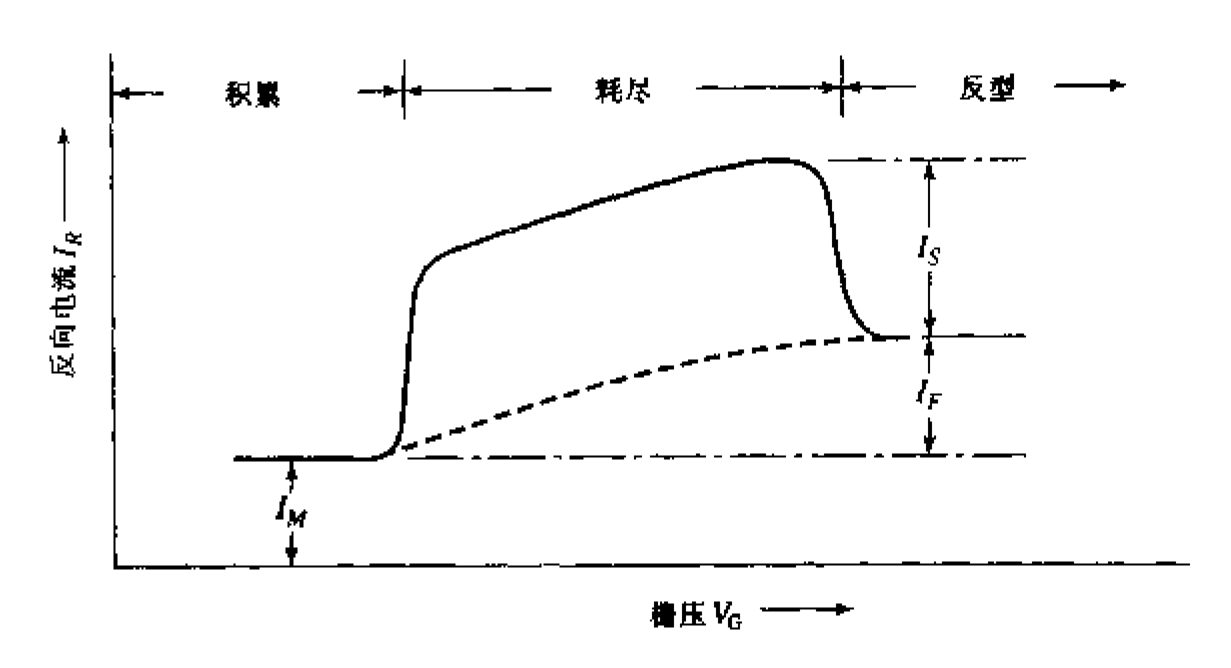


图 8.21　栅控二极管中的反向电流与 V_G的关系曲线，它表明当表面耗尽时漏电流显著增加

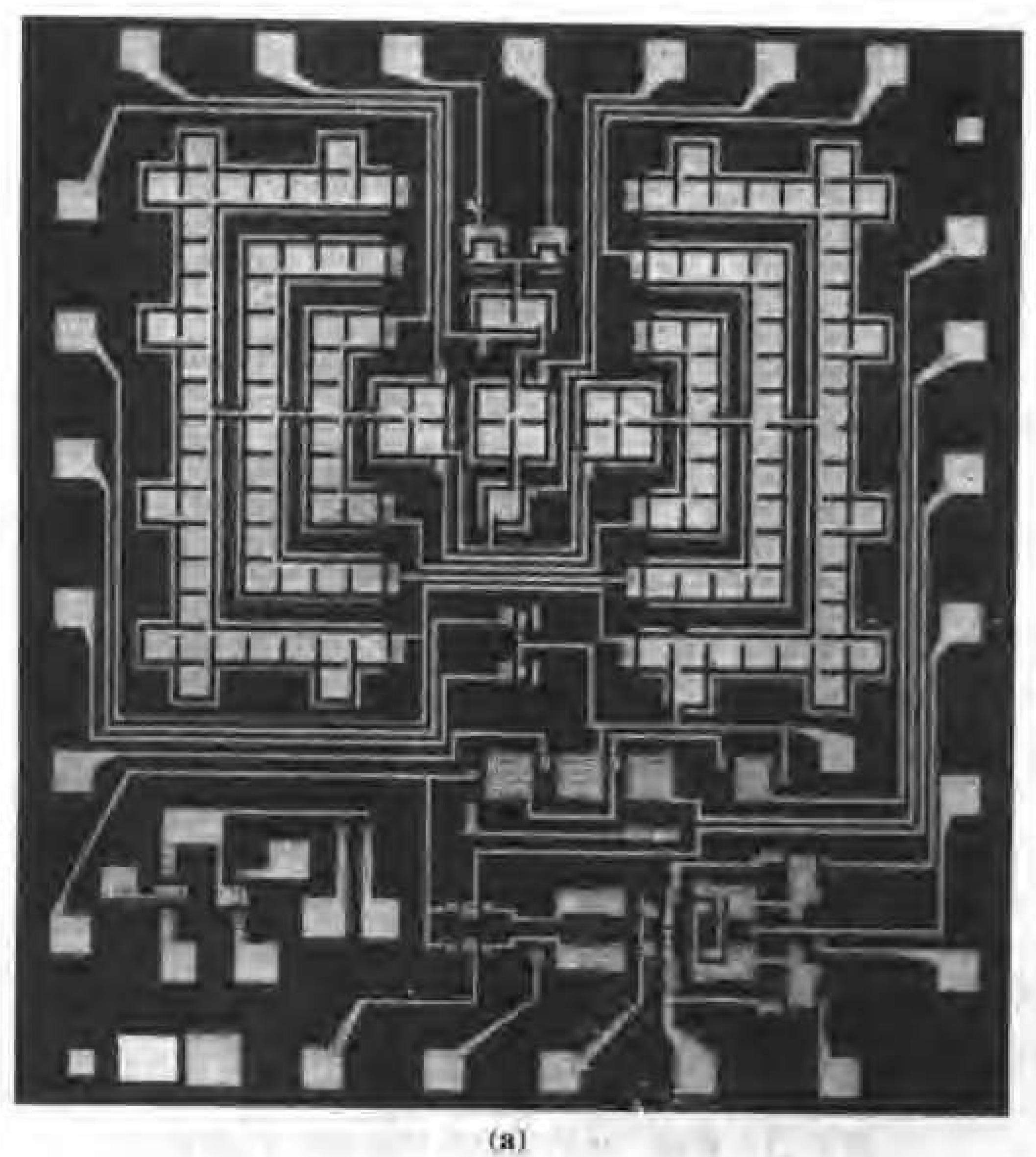
在计算表面电荷层的参数时，必须考虑加在结上的反偏电压与阈值电压的关系。定量地说，利用等式(8.3.18)，可以在阈值电压的计算中考虑结上的反偏电压，此时取沟道电压 V_t 等于二极管的反偏电压 V_R，而衬底电压 V_B为零。

如果不存在栅控二极管结构，则表面状态取决于氧化层电荷。在制造双极型晶体管或 MOS 隔离结构的 IC 工艺中，很多情况下表面电荷在 p 型区会表面产生很宽的耗尽区。此时 pn 结的漏电流很高。通常用退火来减小 s_0和氧化层电荷，从而可显著地改善结的特性。

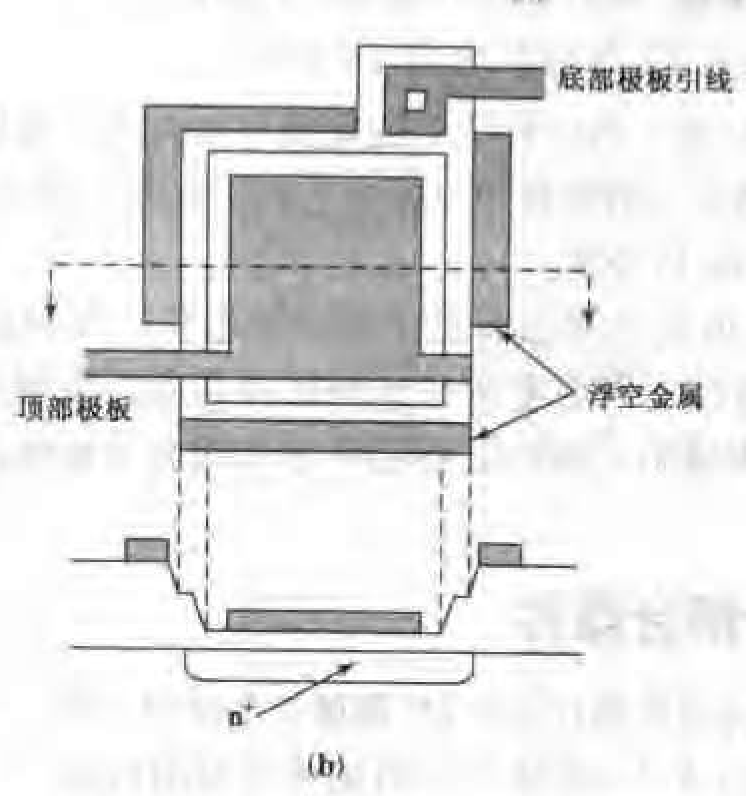
8.7　MOS 电容器及电荷耦合器件

氧化物-硅系统最直接的器件应用是制作高质量、精确控制的电容器。下面举例说明这一应用，图 8.22(a)是一个具有模/数(A/D)转换功能的集成电路的版图[7]。A/D 转换通过顺序地将信号与一个参考电压的各级分压进行比较而完成。此参考电压的分量通过与一系列电容值以 2 的倍数递减的电容上的电压值相比较而实现。

A/D 转换电路中的电容器是 MOS 器件，即图 8.22(a)中的方形结构。为准确地转换模拟信号，这些电容的比值需要精确地加以控制。用图 8.22(b)所示的 MOS 电容就能成功地实现精确控制。图 8.22(b)中浮空的金属条仅用于确保制造电容时刻蚀步骤的精确性，并不具有任何电路功能[7]。



(a)



(b)

图 8.22 (a)具有模数转换功能的集成电路[7]。电路利用了许多如图;(b)所示的精密 MOS 电容器结构

8.7.1 MOS 存储器

在 A/D 转换器中,将电容偏置于积累区可以获得精确的电容值,因为此时 MOS 电容仅对应于氧化层电容,是可精确控制的。

当电容偏置于反型条件时,则可实现另一种完全不同的功能。考虑图 8.23 中所示的 n 沟

道MOS结构，所加正栅压$V_G > V_T$。在稳态时，栅电极下方形成电子反型层（见图8.23(a)），并被一个耗尽层与p型衬底隔开。为屏蔽栅电荷所需的电荷，一部分由负的可动电子提供，另一部分由耗尽层中的电离受主提供，对应于图8.11中的HF曲线。但是我们在8.4节中看到，如果电容是孤立的，则形成反型层的电子需要很长时间的热过程产生。在外加电压的初始阶段，电容处于深耗尽（见图8.23(b)）。所有的用来屏蔽栅电荷的电荷均由耗尽层中的电离受主提供，对应于图8.11中的DD曲线。随着电子的产生，由反型层中可动电子提供的电荷部分开始增加，电容处于深耗尽和反型的中间状态。结果，我们现在得到了这样一个结构，即使在相同的栅压下，包含的可动电荷总量也是可变的，由此我们可以利用存储的可动电荷作为电信号。

由于存储的可动电荷可以代表一种信息，所以MOS电容被广泛地用于存储计算机中的数据。沟道最终被热过程产生的电子填满并达到反型，所以这样的存储单元被称做是动态的。MOS存储单元一般由包含行和列的规则阵列组成，且在两个垂直方向上均可访问某个单元。在这种典型的存储器结构中，可在任意时间同样容易地存取任意单元的内容，所以这种存储结构被称为动态随机存储器或DRAM。

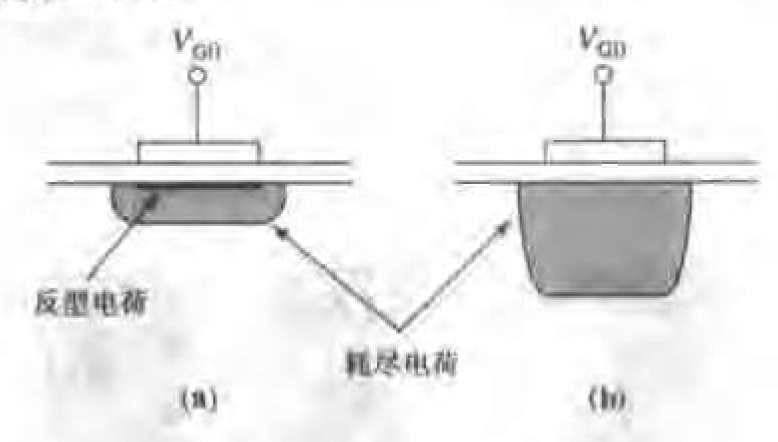


图8.23　具有固定栅极电压V_{GD}的MOS电容器：(a)反型；(b)深耗尽

存储单元一般用于存储数字信号，所以它只有两种状态，即由反型层中产生可动电荷或者缺少可动电荷使电容深耗尽两种状态来表示。电荷的数量由其他器件探测以确定存储的是"0"还是"1"。出于集成度和经济效益的考虑，用于感应和放大电荷信号的电路（灵敏放大器）位于阵列中每列的端部。存储单元首先连接到一条位线（阵列中的竖直列）上，之后被灵敏放大器探测。当存储的电荷与这条位线相连时，就会被位线上的电容泄漏而减少。灵敏放大器上呈现的电压仅仅为$V_{cell} \times C_{cell}/(C_{cell} + C_{bitline})$。随着存储器阵列规模的扩大，位线电容也相应增加。为了正确读出单元的状态，应保持有足够的比值$C_{cell}/(C_{cell} + C_{bitline})$，所以动态存储器的电容应保持足够大，并且不随所用制备工艺的最小特征尺寸的按比例缩小而下降。当最小特征尺寸从2μm减小到250nm时，单元电容的值只是适当减小（以2为因子从50pF减小到~25pF）。为了在芯片上布置更多的存储单元，必须通过减小每个单元的占有面积来提高面积利用率。

在20世纪70年代早期，典型的栅氧化层厚度为200nm。随着栅氧厚度减小到4nm，每单位面积的电容按50倍的比例增大，所以某些低值电容要求面积按2×50=100倍的比例减小。但是，随着绝缘层厚度按50倍比例减小，芯片表面的最小线条特征尺寸也按相同比例缩小（为满足2.1节和第9章中讨论的按比例缩小定律）。结果器件的最小面积按2500倍比例减小，这比电容面积的减小比例100倍要大得多。如果只减小栅氧厚度，那么电容所需的面积就会变得过大，因此在硅衬底面积显著减小的同时，需要采用其他技术仍使电容值能缓慢减小。

电容按$\epsilon_r\epsilon_0 A/x_{ox}$而变化,其中$x_{ox}$是绝缘层的厚度,而$\epsilon_r$是相对介电常数,通常用符号$\kappa$或$K$表示。可以改变表达式中的一些因子以增加电容值与特征尺寸的比值。(1)随着工艺的发展,绝缘层的厚度可以减小。其下限要保证漏电流足够小,并且可形成面积均匀、无缺陷的绝缘层。纯SiO_2的极限厚度约为2~3nm。(2)可以使用高介电常数的绝缘材料。例如ϵ_r为7的氮化硅,有时与氧化硅(ϵ_r~4)结合起来使用。也可以考虑使用更高介电常数的材料如五氧化钽(Ta_2O_5,ϵ_r~22)和更复杂的合金材料,如ϵ_r高达几百的钡和钛酸锶的合金($Ba_{1-x}Sr_xTiO_3$)。(3)在第三维上扩展单元或者用不规则的、有起伏的表面来增加表面积(从而增加电容),这些方法都可以使芯片表面的单元占有面积得以更充分的利用。

在某些情形,在衬底中刻蚀出沟槽,其侧壁被氧化形成栅绝缘层,沟槽内被电极极板填充,如图8.24(a)所示。沟槽电容器的有效面积大约为$4F\times D$,其中F是表面特征尺寸(最小的表面线条尺寸),而D是沟槽的有效深度。此时,存储电荷的反型层从垂直方向向邻近的衬底区扩展。尽管这一技术显著地增加了有效面积,但是也很快遇到了极限。当相邻电容相隔太近时,分隔相邻电容的中性区被"穿通",侧向包围电容的耗尽区因而相互影响(见图8.24(b))。相邻电容间的最小间隔为$2x_d$,其中x_d是最大耗尽区宽度,这实际上对应于深耗尽状态(即电容上没有存储可动电荷)。另一种技术是所有单元共用置于阱外的电极极板,并保持固定电压。这时压控电荷存储区位于沟槽内;这个区域为高掺杂,所以电容并不随电压显著地变化。

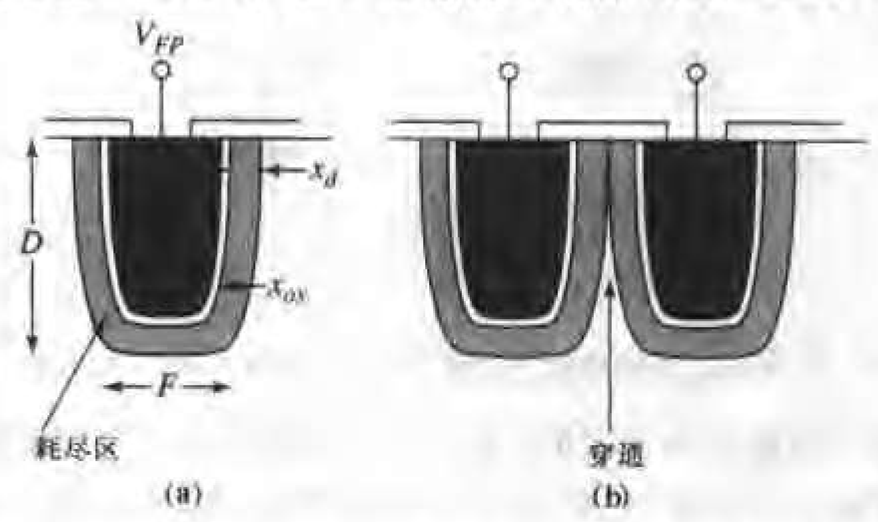


图8.24 (a)特征尺寸为F,阱深为D的沟槽型电容器的截面图;(b)当耗尽区相互接触时两个相邻的沟槽型电容器相互作用

8.7.2 电荷耦合器件

上面描述的MOS电容用于存储数字信号,所以只需要两种状态。同样,模拟信号也可以存储在MOS电容上。根据上面讨论的偏置情况,工作于深耗尽和反型之间的MOS电容是电荷耦合器件(CCD)[8]的基础,它不仅具有存储功能,还可以转移电荷。电荷的转移是通过距离足够近的电容阵列来实现的,这样MOS电容中的反型层(沟道)中的可动电荷在适当的电极偏置下可转移到邻近电容的沟道中。

CCD被广泛应用于光学成像系统以探测入射光。由于可以形成高密度探测器阵列,因而可获得高分辨率的图像。在这一应用中,CCD单元或图像元素(像素)被偏置在深耗尽状态,并且在聚焦图像下曝光一段时间。进入硅体内的光子以一定的速率产生出电子-空穴对,其产生率与图像的亮度成正例。每个单元的耗尽区中的电场将电子和空穴分离,将电子推向氧化物-硅界面存储起来。于是每个栅下的沟道开始带电,其荷电的程度代表了该位置上图像的亮

度。随着光照的继续，表面势最终接近于反型值，且存储的电荷数量趋于饱和。因此，单元只需被光照有限的时间。电荷产生后，单元中的模拟信息沿着 CCD 成像阵列进入外围的探测电路系统。这一设计用于集成电路静态摄像和电视摄像。

CCD 的工作速度受到热产生过程的限制。光信号激发的电荷必须以一定的速度被探测并移出阵列，这一速度比硅表面耗尽区本身电荷的产生速度要快。任何 CCD 单元都不能在深耗尽条件下保持很长的时间，因为如果热产生过程的时间超过了信号经过 CCD 阵列的周期时间，会引入噪声。

这些限制使得对氧化物-硅系统的质量有非常严格的要求。如果表面产生-复合中心的浓度能控制在非常低的水平，这些要求才可以满足。CCD 电路的制造者们一般将这些态密度值控制在 $10^9 cm^{-2}$ 的范围内，以使表面产生流处于 $1 \sim 10$ nA cm^{-2} 的范围。事实上很多 CCD 电路中只有很少数量的电荷。由于 CCD 栅电极面积的典型值为几平方微米，所以可探测电荷的值大致从 10 个电子到 10^7 个电子。相对而言，$1(\mu m)^2$ 的栅下 1nA cm^{-2} 的热产生电流对应的电荷大约为每秒 60 电子。

例题　电荷耦合器件

我们想设计一个电荷耦合器件作为图像传感器，图像元素（像素）由边长为 $5\mu m$ 的正方形栅电极组成。可探测的电荷阈值为 2500 电子/像素，对每一像素上的电荷采样，且每 10ms 复位清零一次。热平衡时，CCD 的反型层的电荷密度为每平方厘米 10^{13} 个电子。

假如电子的热（无光照）产生率与时间成指数关系（见习题 8.6 中的推导），在 $12\Omega \cdot cm$ 的 p 型硅中，如果可探测电荷阈值中少于 5% 的电子由热过程产生时，试求所对应的少数载流子寿命 τ_0。

解：热平衡时，每个栅下有 $10^{13} \times (5 \times 10^{-4})^2 = 2.5 \times 10^6$ 个电子，可探测电荷阈值为 2500 电子，每个栅上允许由热过程产生电子的数量是 $2500 \times 0.05 = 125$。

因此 $2.5 \times 10^6[1 - \exp(-t/\tau_a)] = 125$，所以 $t/\tau_a = 5 \times 10^{-5}$。在产生时间 $t = 10^{-2}$ s 内，满足要求的最小表面产生寿命 τ_a 为

$$\tau_a = \frac{10^{-2}}{5 \times 10^{-5}} = 2 \times 10^2 s$$

根据习题 8.6 的分析

$$\tau_0 = \frac{n_i}{2N_a} \times \tau_a$$

从图 1.15，电导率为 $12\Omega \cdot cm$ 的硅，所对应的 $N_a = 10^{15} cm^{-3}$；因此最小可接受寿命为

$$\tau_0 = \frac{1.45 \times 10^{10}}{2 \times 10^{15}} \times 2 \times 10^2 = 1.45 \times 10^{-3} s = 1.45 \text{ ms}$$

这个习题的结论强调了在 CCD 中表面产生特征时间 τ_a 比少数载流子寿命 τ_0（1.45ms）长得多（200 s）。

为了讨论 CCD 中的电荷转移，假设电荷的产生和迁移时间均比反型层热产生过程所需时

间短。考虑图 8.25 所示的 MOS 系统,图中三个电容并列于硅表面。假设中间栅上所加的电压 V_2高于两边的栅压,并高于 MOS 的阈值电压。如果电子被引入表面区,它们将滞留在中间栅下的沟道内。如果在右边的栅上加更高的电压 V_3,将引起电子向该栅下的沟道转移(图 8.25(b))。此时中间栅上的电压降到 V_1,接着右边栅上的电压减小到 V_2。最后的结果是沟道中的电荷转移到右边的栅下。

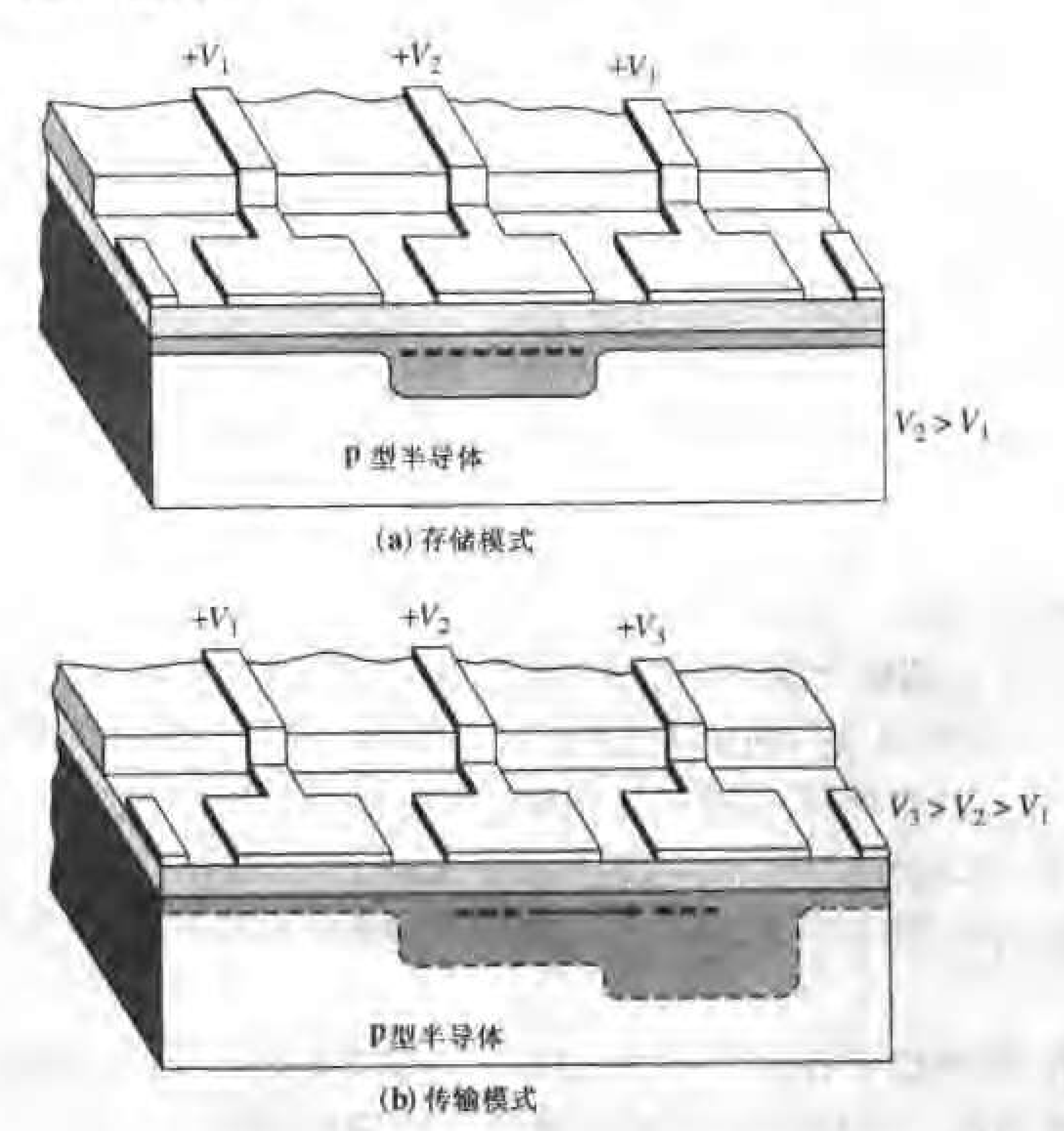


图 8.25 CCD 中的基本传输原理:(a)存储模式,电荷位于中间栅下的沟道中;(b)右边栅上的电压 $V_3 > V_2$时引起电荷向右边转移

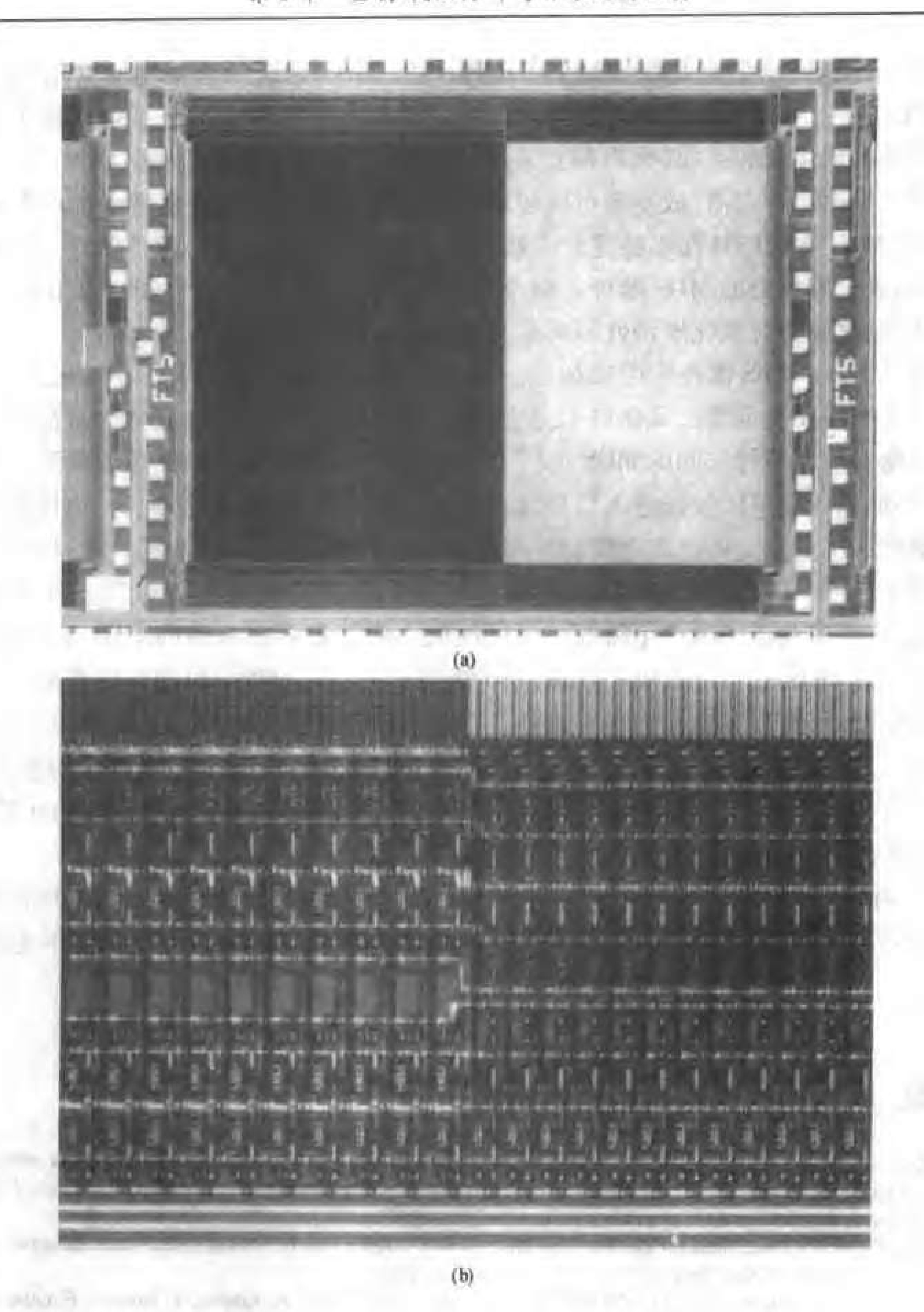
CCD 制造中的一个关键因素是相邻栅之间的间隔。为使电荷在边缘电场的作用下实现预期的转移,栅必须靠得足够近。尽管过去的系统的构成仅利用了金属栅,但是单层硅栅或两层或三层硅栅也很有用,尤其是在设计精细的时钟电路时。

图 8.26 为高密度、帧变换成像 CCD 的一部分。这个 CCD 使用三层多晶硅和一个具有 p 阱的 n 型衬底。控制 CCD 的外围电路由 CMOS 工艺实现(见第 9 章)。

除了成像,CCD 的另一个应用是作为延时线。通过在不同的延时周期读取信号,各种有用的信号处理技术,例如卷积(涉及到时变信号与其延时后的信号相乘)可很容易地用 CCD 实现。

CCD 也可以用于存储数字信号。由于仅有两种状态需要被存储("0"和"1"),所以在相同数量级下对噪声和电荷产生的要求要松一些。但是,需要读取的位(以及相同列的所有位)必需用时钟信号选中,并将信息传输到芯片末端的探测电路。由于获取任意位的信息的时间周期不稳定,且有时很长,这使得该类存储器不如随机存储器灵活。

从这些讨论中可以看出,CCD 是非常有用的成像器件,也适用于信号处理和数字存储。



(a)

(b)

图8.26　(a)高密度帧变换CCD。探测阵列位于芯片的左边，光屏蔽的存储区在右侧。该CCD的像素为588×605，芯片面积为38.22mm²；(b)CCD周边的CMOS电路。(1984,IEEE)[10]

要在氧化物-硅界面形成CCD，要求热载流子的产生率很低，因此需要对这一工艺以及MOS系统的电学特性有详细的了解。

小结

通过考虑外加电压对金属-氧化物-硅(MOS)系统的作用，详细研究了氧化物-硅系统的电

学特性。MOS 结构的一个重要参量是平带电压 V_{FB}。MOS 系统的能带图非常清楚地说明了 V_{FB}的重要性。当金属和衬底间加电压 V_{FB}时,硅表面没有电场和电荷。平带电压由几个参量决定。对理想的氧化层和界面(无电荷),V_{FB}仅取决于金属和半导体的功函数差。但是实际结构中氧化层和界面均有电荷,故平带电压还取决于这些电荷。如果氧化层中的电荷量随外加电压变化,则 MOS 系统的特性不稳定。一般来说,氧化层电荷密度必须控制在 $10^{10}cm^{-2}$ 量级以下,才能制造出可接受的 MOS 器件。碱金属,尤其是钠,可以在加于 MOS 氧化层上的电场作用下移动,因此成为对氧化层特性影响最大的杂质。

很多情况下,在 MOS 器件中用掺杂的硅制作"金属"电极非常有用。这种硅是用化学气相淀积的方法生长在无定型二氧化硅上,是由很多小晶粒组成的,因此被称为多晶硅。

在金属栅或硅栅与衬底间的 MOS 电容上加电压,可以使硅表面处的多数载流子积累或耗尽。同样,外加偏压也可以使硅进入反型状态而在硅的表面引入 pn 结,此时,由栅控制的电荷中一部分是耗尽层中固定的电离杂质,另一部分是反型层中的自由载流子。通过测量 MOS 电容可测定积累、耗尽和反型这三种状态。测量得到的反型区 MOS 电容取决于外加电压的频率,因为形成反型层载流子的产生需要一定的时间。对于有界面陷阱的实际的 MOS 系统,测量得到的耗尽区 MOS 电容也取决于频率,因为陷阱的充放电过程与时间常数有关。

当表面反型时,反型层通常被称为沟道。如果沟道存在于 pn 结附近,该感应 pn 结可以通过一个扩散 pn 结来加偏置。这种偏置方法可以改变耗尽层和反型层中栅感应的电荷分布,其中耗尽层电荷是固定的,反型层电荷是可动的。硅表面的空间电荷区可以影响 pn 结的特性。与栅控二极管的特性一样,表面态的行为也取决于表面势。

MOS 结构的一个重要而直接的应用是制备集成电路电容器,这些电容通常用在动态随机存取存储器和模数转换器中。排列很近的 MOS 电容阵列可用于形成光学成像的电荷耦合器件(CCD)。

参考文献

1. R. H. KINGSTON and S. F. NEUSTADTER, *J. Appl. Phys.* **26**, 718 (1955).
2. C. E. YOUNG, *J. Appl. Phys.* **32**, 329 (1961).
3. S. M. SZE, *Physics of Semiconductor Devices, 2nd Edition*, Wiley-Interscience, New York, 1981, p. 372.
4. B. E. DEAL, *J. Electrochem. Soc.* **127**, 979 (1980).
5. C. N. BERGLUND, *IEEE Trans. Electron Devices*, **ED-13**, 701 (1966).
6. A. S. GROVE and D. J. FITZGERALD, *Solid-State Electron.* **9**, 783 (1966).
7. J. L. MCCREARY and P. R. GRAY, *IEEE J. Solid-State Circuits*, **SC-10**, 371 (1975). Reprinted by permission.
8. W. S. BOYLE and G. E. SMITH, *Bell Sys. Tech. J.* **49**, 587 (1970).
9. C. JUND and R. POIRER, *Solid-State Electron.* **9**, 315 (1966).
10. N. TERANISHI, A. KOHNO, Y. ISIHARA, E. ODA, and K. ARAI, *IEEE Trans. Electr. Devices*, **ED-31**, 1829 (1984).
11. A. J. P. THEUWISSEN, C. H. L. WEIJTENS, L. J. M. ESSER, J. N. G. COX, H. T. A. R. DUYVELAR, and W. C. KEUR, *Tech. Digest IEEE Int. Electr. Devices Mtg.* (Dec. 1984), p. 40.

习题

8.1* 画出理想的铝栅 MOS 系统在(i)热平衡情况,(ii)平带情况下的能带图,已知(a)1Ω·cm 的 n 型硅和(b)1Ω·cm 的 p 型硅。

8.2 将 8.1 题中 MOS 系统的铝栅改为多晶硅栅,重复完成 8.1 题所要求的能带图。假设硅栅与单晶硅有类似的能带结构,但是

(i) n 型硅上的栅掺入受主杂质，刚刚达到简并；

(ii) p 型硅上的栅掺入施主杂质，刚刚达到简并。

(这些条件对应于常用的硅栅技术，原因将在第 9 章中解释)

8.3　证明等式(8.4.4)给出的偏置于耗尽状态下的小信号 MOS 电容 C。也就是说，证明 C 等于以下两个电容的串联：(1)由硅体为一个极板，氧化物-硅界面为另一个极板构成的电容；(2)被氧化层分开的两极板构成的电容。

(提示：利用 Gauss 定理表示电荷 $\Delta Q = \epsilon_{ox}\Delta\mathscr{E}_{ox}$，然后证明电容两端的电压 $\Delta V = \Delta\mathscr{E}_{ox}x_{ox} + \Delta\mathscr{E}_{ox}\epsilon_{ox}x_d/\epsilon_s$，并计算 $C = \Delta Q/\Delta V$。)

8.4　已知 $V_{FB} = -0.5\text{V}$，利用等式(8.4.4)求出 MOS 系统处于耗尽区时的总电容 C 的表达式。画出 C/C_{ox} 对 V_G 变化的曲线。考虑二氧化硅层厚 100nm，衬底是 $1\Omega\cdot\text{cm}$ 的 p 型硅，利用等式(8.4.3)求平带电容 C_{FB}。

8.5　求 MOS 系统的实验研究时经常要用到表面势 ϕ_s 值。

(a) 利用习题 8.3 的结论，证明处于耗尽区的 MOS 电容，当其栅压 V_G 变化时，通过测量得到的电容值可以计算出相应的表面势 ϕ_s 的变化。ϕ_s 的变化可以由以下关系式算出

$$\phi_s(V_{G2}) - \phi_s(V_{G1}) = \int_{V_{G1}}^{V_{G2}}\left(1 - \frac{C}{C_{ox}}\right)dV_G$$

这一方法以发明者的名字命名[5]为 Berglund 方法。如果取 $V_{G1} = V_{FB}$，利用等式(8.4.3)可非常方便求出该点的 C 值。

(b) 如果取 $V_{G1} = V_{FB}$，画出 p 型硅的低频 MOS 电容曲线(用 C_{ox} 归一)，并在曲线上用阴影表示出等于 $\Delta\phi_s$ 的区域。

8.6[†]　考虑 p 型硅上的 MOS 系统，$t = 0$ 时刻在栅上突然加总量等于 Q_G 的电荷，使其偏置于深耗尽区根据等式(5.3.26)，硅表面空间电荷区由载流子产生引起的沟道充电电流可用沟道电荷 Q_n 表示为

$$\frac{dQ_n}{dt} = -\frac{qn_i(x_d - x_{df})}{2\tau_0}$$

式中 x_d 是表面耗尽区宽度(与时间有关)，τ_0 是等式(5.2.14)给出的电子寿命。x_{df} 是热平衡下的空间电荷区宽度；也就是说，当 $x_d = x_{df}$ 时，由于载流子产生而引起的沟道充电电流为 0。

(a) 证明 Q_n 的微分方程为

$$Q_n + \left(\frac{2\tau_0 N_a}{n_i}\right)\left(\frac{dQ_n}{dt}\right) = -[Q_G - qN_a x_{df}]$$

(b) 在 $Q_n = 0(t = 0)$ 的条件下求解该方程。并说明形成表面反型层的特征时间为 $2N_a\tau_0/n_i$ 量级[10]。

8.7　画出如图 P8.7(a)，图 P8.7(b)，图 P8.7(c) 所示 MOS 结构的电容-电压曲线。测量频率为 100kHz，并用氧化层电容将这个小信号电容归一化。测量时栅上的直流偏置变化缓慢。用虚线说明正电荷 Q_f 的增加对 $C\text{-}V_G$ 曲线产生的影响，并在曲线上注明各个区(积累，耗尽，反型)。假设衬底电阻率为 $10\Omega\cdot\text{cm}$ 量级，只要求定性画出正确的曲线。

8.8　如果 $V_{FB} = -2\text{V}$，并且主要是氧化层固定电荷 Q_f 的作用。画出偏置在反型时 n 型衬底的 MOS 电容曲线。图中应包括(a)能带图(b)系统中的所有电荷(c)电场分布(d)电势分布(以硅体内作为电势参考点)

8.9　建立一个描述 n 型硅表面电荷状态的类似于表 8.1 的表格。

8.10　利用 8.3 节中的公式，证明当 $n_s = 10N_a$ 时，ϕ_s 仅比 $-\phi_p$ 高 58mV。

8.11[†]　考虑 Q_n(等式(8.3.16))和 V_T(等式(8.3.18))与 $(V_G - V_B)$ 的依赖关系，定性绘出一组当 $(V_G - V_B)$ 变化时，C/C_{ox} 与 V_G 的关系曲线(低频)。这个关系由 Grove 和 Fitzgerald 研究得到[6]。

8.12*　计算(a) $1\Omega\cdot\text{cm}$ p 型硅，以及(b) $1\Omega\cdot\text{cm}$ n 型硅的阈值电压。两种情况下的 MOS 系统为(i) $q\Phi_M = 4.1\text{eV}$ 的铝栅(ii) 100nm 厚的二氧化硅层，(iii)二氧化层表面只含有面密度为 $(Q_f/q) = 5\times10^{10}\text{cm}^{-2}$ 的

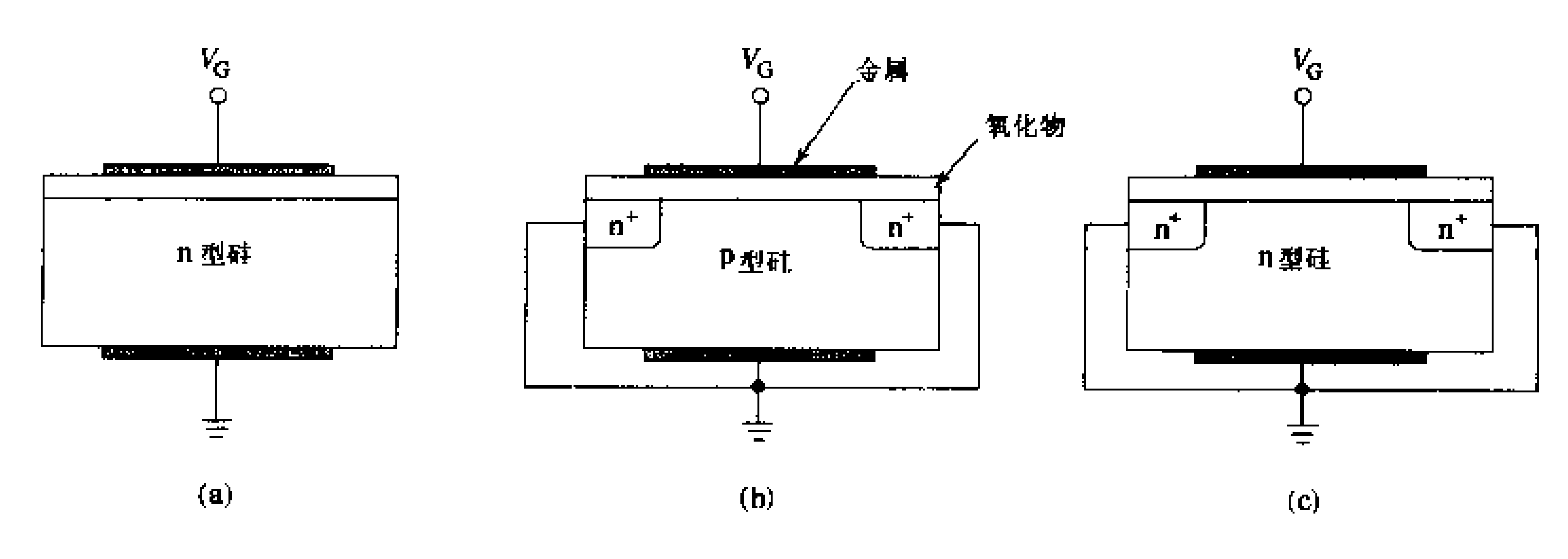


图 P8.7

电荷。除栅压之外沟道无其他偏置($V_C = V_B = 0$)。

8.13* 考虑8.5节描述的表面氧化层电荷对n型区或p型区的不同影响,分析双扩散双极型晶体管的高电阻率集电区的情况。用图形和讨论说明,为什么这些效应使制造稳定的可重复的双扩散pnp双极晶体管比制造npn晶体管困难。

8.14† 在实际MOS系统中,C-V测量有时呈现迟滞效应。也就是说,C-V_G曲线有图P8.14所示的形状。这条曲线对应的测量条件是V_G用低频三角波(~1Hz)扫描,交流测量频率在1kHz量级或者更高。不论如图所示的逆时针测量方向还是顺时针测量方向,均可在曲线上观测到迟滞现象。利用这一现象可以区分导致非理想特性的两种最常见的原因。

(a)考虑以下非理想效应说明:

(i)在电场作用下绝缘层中正离子的运动;

(ii)氧化物-硅界面陷阱Q_{it}对沟道自由载流子的陷阱作用。

(b)用图形定性说明n型和p型衬底的C-V_G曲线;在每幅图上标明迟滞现象(顺时针或逆时针)

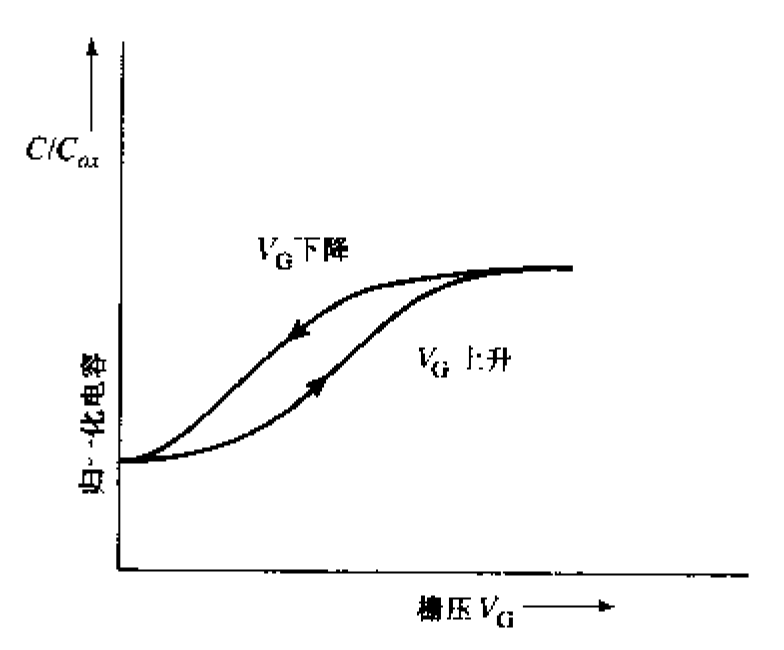


图 P8.14

8.15* 已知电容面积为$100\times100\mu m^2$,试比较MOS电容或反偏pn结二极管可得到的最大电容值。假设MOS氧化层的击穿场强为$8\times10^6 V\cdot cm^{-1}$,工作电压为5V,安全系数为2(即设计的MOS氧化层可承受10V的电压)。pn结为在n型硅中掺入$10^{16}cm^{-3}$的硼形成。

8.16† 当表面完全耗尽时由表面态导致的产生流I_S(等式(8.6.3))等于表面耗尽区的产生流I_r(等式8.6.2)的两倍时,计算表面态面密度。已知表面态的俘获截面为$10^{-15}cm^2$,热运动速度为$10^7 s^{-1}$,表面耗尽区宽度为$1\mu m$,时间常数τ_0为$1\mu s$。

8.17 测量如图P8.17(a)中所示的两相邻的栅与衬底之间的MOS电容C_1和C_2时,得到相同的C-V特

性(图 8.17(b)中的曲线由耗尽近似得出)。当衬底浮空时,在测量两栅之间的电容时,画出得到的电容与电压的关系曲线,并在图线上标出重要的特征参数。

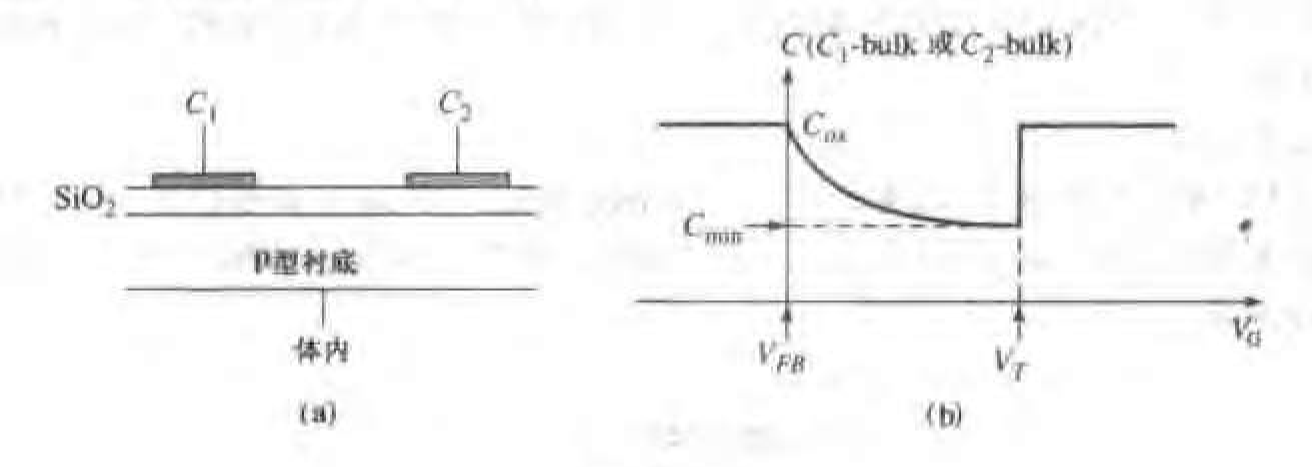


图 P8.17

8.18　如图 P8.18(a)中所示的电容 C_A 是理想 MOS 电容,其 Si-SiO_2 界面没有陷阱和氧化层电荷。电容 C_B 和 C_C 的尺寸,衬底导电类型及掺杂浓度均与 C_A 相同,但包含 $10^{11}\,cm^{-2}$ 的氧化层电荷密度和如图 P8.18(b)

(a)

C_B 和 C_C 的界面陷阱分布

(b)

(c)

图 P8.18

所示的界面陷阱分布。P8.18(c)给出了平衡条件下 C_A 的 C-V 曲线。已知 $x_{ox}=100nm$ 且 $N_a=10^{15}cm^{-3}$。

(a)画出 C_B 的高频 C-V 曲线(陷阱跟不上交流信号)和低频 C-V 曲线(所有陷阱都跟得上交流信号),将它们叠加在图 P8.18(c)所示的 C_A 的 C-V 曲线之上,并在曲线中标出重要的特征参数。假设 MOS 系统在直流偏置下处于平衡。

(b)对 C_C 重复(a)。

8.19 图 P8.19 为 MOS 栅控二极管的 C-V 曲线的测量装置。栅极电压 V_G 是扫描速率为 R_m V/s 且很慢的从正到负的线性斜坡电压,端电流 I_1,I_2 和 I_3 由三个灵敏度非常高的静电计监测。在同一个图中画出 I_1,I_2 和 I_3 与 V_G 的关系曲线,并解释该图。

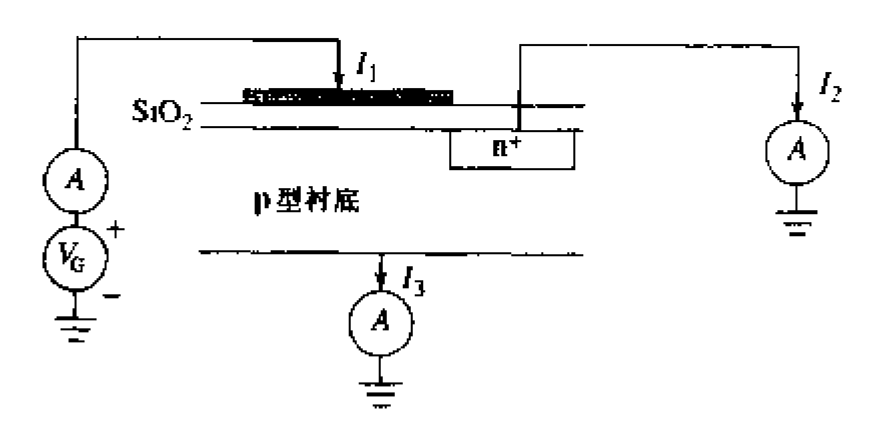


图 P8.19

表 8.3 氧化物-硅系统的常用公式

p 型衬底(n 沟)	n 型衬底(p 沟)
平带电压 (等式(8.5.6)) $V_{FB}=\Phi_{MS}-\dfrac{Q_f}{C_{ox}}-\dfrac{1}{C_{ox}}\int_0^{ox}\dfrac{x}{x_{ox}}\rho(x)dx$	
体内电势(等式(4.2.9)) $\phi_p=-\dfrac{kT}{q}\ln\left(\dfrac{N_a}{n_i}\right)$	$\phi_n=\dfrac{kT}{q}\ln\left(\dfrac{N_d}{n_i}\right)$
强反型表面势(表 3.1) 热平衡 $\phi_s=\|\phi_p\|$ $\phi_s-\phi_p=2\|\phi_p\|$	$\phi_s=-\|\phi_n\|$ $\phi_s-\phi_n=-2\|\phi_n\|$
偏压 $(V_C-V_B)=V_{CB}$ $\phi_s=\|\phi_p\|+V_{CB}$	$\phi_s=-\|\phi_n\|-\|V_{CB}\|$
最大耗尽层宽度 x_{dmax}(等式(8.3.6)),热平衡 $\sqrt{\dfrac{4\epsilon_s\|\phi_p\|}{qN_a}}$	$\sqrt{\dfrac{4\epsilon_s\|\phi_n\|}{qN_d}}$
偏压 V_{CB}(等式(8.3.8)) $\sqrt{\dfrac{2\epsilon_s(2\|\phi_p\|+V_{CB})}{qN_a}}$	$\sqrt{\dfrac{2\epsilon_s(2\|\phi_n\|+V_{CB})}{qN_d}}$
功函数差 Φ_{MS} $\Phi_M-(X+E_g/2q+\|\phi_p\|)$	$\Phi_M-(X+E_g/2q-\|\phi_n\|)$
阈值电压(任意参考点)(等式(8.3.18)) $V_{FB}+V_C+2\|\phi_p\|+\dfrac{1}{C_{ox}}\sqrt{2\epsilon_s qN_a(2\|\phi_p\|+V_C-V_B)}$	$V_{FB}+V_C-2\|\phi_n\|-\dfrac{1}{C_{ox}}\sqrt{2\epsilon_s qN_d(2\|\phi_n\|+V_B-V_C)}$

第9章 MOS场效应晶体管Ⅰ:物理效应和模型

实际上在双极晶体管发明之前,就已经提出了金属-氧化物-半导体晶体管的概念。20世纪30年代早期,就出现了类似现代的硅MOSFET的专利,但是这些MOSFET用的是不含硅的化合物材料[1]。当时不能制造出现代的硅MOSFET晶体管的原因主要有两个:对于绝缘体-半导体系统的物理特性的认识还不完善;难以控制绝缘体和半导体之间的界面特性。然而,相关的专利文献表明,在随后的几十年里通过人们的不断努力,才有了现代的MOSFET晶体管。

William Shockley的固态电子学经典论文"半导体中的电子和空穴"[1]描述了20世纪40年代ATT贝尔实验室在这方面所作的艰苦努力。当时他们的器件类似于一个平板电容器,其中的一极为半导体,另一极为靠得很近的金属。"当金属上施加正电压的时候,半导体中出现额外增加的电子电荷,这些电子可以自由移动,并对半导体导电作出贡献。"贝尔实验室的努力首先导致在1948年发明了双极晶体管,不久之后就诞生了第一个可以工作的场效应晶体管。

但是,第一个商用MOSFET器件直到60年代才出现。以后的四十多年里,MOSFET代替了双极晶体管成为最广泛使用的有源固态器件。与双极晶体管相比,MOSFET在数字电路方面有着非常多的优点,而在同一时期,数字电路也呈指数发展趋势。MOSFET的工艺要求比双极晶体管简单,并有利于集成,而且MOSFET集成电路的功耗比双极晶体管集成电路小,特别是在低频工作的时候。基于上述特点(工艺简单、密度高、功耗低),MOSFET在存储器电路中有着广泛的应用,完全取代了双极晶体管存储电路。这些优点也导致了MOSFET在逻辑电路中的巨大优势,尤其是高速微处理器(本书封面和图2.2(e))。MOSFET集成电路的大批量生产,使得有更多的资金投入到为提高MOSFET的性能而进行的研究中,其中器件的特征尺寸一直在不断地减小,使得单片上可以集成更多的器件,并提高了频率响应特性。经过30多年的发展,MOSFET的特征尺寸已经下降了200倍,由原来的20μm缩小到0.1μm以下,相应的单位面积可以集成的器件数也提高为原来的4万倍。

按比例缩小技术指的是为提高集成度和电路的工作速度,同时缩小有源器件和无源器件的尺寸。1974年,Dennard等人[2]在按比例缩小和电路性能的关系方面做出了有影响的系统研究,并提出了恒场缩小理论。按照该理论,随着器件尺寸的缩小,晶体管内部的电场应尽可能地保持不变。表9.1总结了恒场缩小规律。

实际上该理论不能被完全严格地执行,所以这一理论只是对集成电路的尺寸缩小起指导作用。决定有源和无源器件尺寸的更重要的因素是工艺和器件性能的限制,而不是仅仅依靠几何尺寸的按比例缩小。表9.2列举出了过去的和未来预测的共七代MOS按比例缩小趋势,

---

[1] 早期的专利包括:J. E. Lilienfeld,美国专利,1,745,175(1930年1月28日),1,877,140(1932年9月3日),1,900,018(1933年3月7日);O. Heil,英国专利,439,457(1935年12月6日)。

表9.1 恒场按比例缩小规律(© 1974年 IEEE[2])

物理参数	按比例缩小因子
表面尺寸,L	$1/K$
纵向尺寸,x_{ox},x_j	$1/K$
掺杂浓度	K
电流,电压	$1/K$
电流密度	K
电容(单位面积)	K
跨导	1
电路延迟时间	$1/K$
功耗	$1/K^2$
功耗密度	1
功耗-延迟积	$1/K^3$

这些数据由ITRS[3]提供。从表中的数据我们发现,金属连线的宽度和厚度缩小的比例远远不如MOSFET器件的沟道长度L的缩小比例,原因有两点:在不平的表面光刻小尺寸、高反射率的金属线条比较困难;由于电迁移效应导致电流密度不能过大(见2.7节)。Dennard提出的恒场律存在着固有的局限性:在大尺寸器件中可以被忽略的效应,随着器件特征尺寸的缩小而变得不可忽略,并随着尺寸的进一步缩小而成为主导因素,因此恒场律的使用受到限制。MOSFET的沟道长度目前已经达到深亚微米量级,此时器件的性能主要由强场效应来决定的,这与大尺寸器件完全不同。本章将讨论MOSFET最基本的物理模型和短沟器件的物理模型。基本物理模型通过分析大尺寸器件来建立解析模型,并从物理上了解器件基本的工作原理,接着将考虑尺寸缩小引起的附加效应,并对模型进行相应的修正。

表9.2 过去的和未来预测的MOS技术的按比例缩小趋势,引自ITRS(1997-2001年版本[3])

生产年份	1997	1999	2001	2003	2006	2010	2016
最小尺寸L(μm)	0.25	0.18	0.13	0.1	0.07	0.045	0.022
DRAM密度(Gbits/cm^2)	0.18	0.38	0.42	0.91	1.85	4.75	28.85
电源电压V_{DD}(V)	2.5-1.8	1.8-1.5	1.2	1.0	0.9	0.6	0.4
等效x_{ox}(nm)	4-5	1.9-2.5	2.3	2.0	1.9	1.2	0.9
结深x_j(nm)	50-100	45-70	30-60	26-52	20-40	15-30	10-20
局部连线跨距(nm)	600	500	350	245	130	105	30
金属高宽比	1.8	2	1.6 *	1.6	1.7	1.8	2.0
max. I_{Dsat} NMOS	600	750	900	900	900	900	900
(μA/μm) PMOS	280	350	420	420	420	420	420

* 改为铜布线, 可用于生产的工艺技术已经成熟

9.1节讨论MOSFET最基本的工作原理,由此推出"一阶"模型。9.2节给出考虑了迁移

率下降和漂移速度饱和这两个最重要的二阶效应的改进模型。9.3节分析CMOS电路和电路的设计方法,并讨论如何得到所需要的参数。9.4节进一步分析MOSFET设计未来的发展方向。MOSFET沟道区的强场效应是器件设计中需要足够重视的因素,我们将在第10章中深入分析。尽管考虑到细节问题,MOSFET的器件分析会引入越来越多的数学公式,我们还是希望尽可能给大家提供一个清晰的物理图像,而不是单纯的列举枯燥无味的公式。

9.1 MOSFET的基本工作原理

MOS晶体管分为n沟道(导电载流子为电子)和p沟道(导电载流子为空穴)两种类型,通常,我们将这两种晶体管称为NMOSFET和PMOSFET。典型的n沟道MOSFET做在p型硅衬底上,这样可以利用反偏的pn结来隔离相邻器件的导电沟道;相应的,p沟道MOSFET做在n型衬底上。对于n沟道MOSFET,如果在栅上加足够高的正电压,就会在栅下形成导电沟道;p沟道MOSFET在栅上则要加足够高的负电压。由于n沟道MOSFET非常重要,下面我们将只分析n沟道MOSFET,得出的结论同样适用于p沟道MOSFET,只需要改变某些参数的正、负号即可。

图9.1为MOSFET的基本结构:两个n型区分别是源区和漏区,它们之间的横向距离就是沟道长度L。通常规定把指向硅衬底、垂直于氧化层-硅界面的方向定义为x轴(如第8章),沿沟道方向定义为y轴。一般源区和漏区是断开的,只有当硅表面形成了n型反型层时,才有导电沟道将源漏相连。当表面反型,并且在源极和漏极间上加上电压时,载流子将从源极流入沟道,从漏极流出,形成电流。考虑到导电载流子的极性,n沟道MOSFET的电流正方向是从漏极到源极,p沟道MOSFET是从源极到漏极。另外,如果栅源偏压为零时,MOSFET中就已经存在反型沟道,这时也可以通过控制栅源电压来减小漏源电流,这种类型的晶体管称为耗尽型晶体管。而前面提到的在$V_{GS}=0$时不存在反型沟道的晶体管称为增强型晶体管。增强型晶体管在电路中的使用比耗尽型晶体管要多一些。

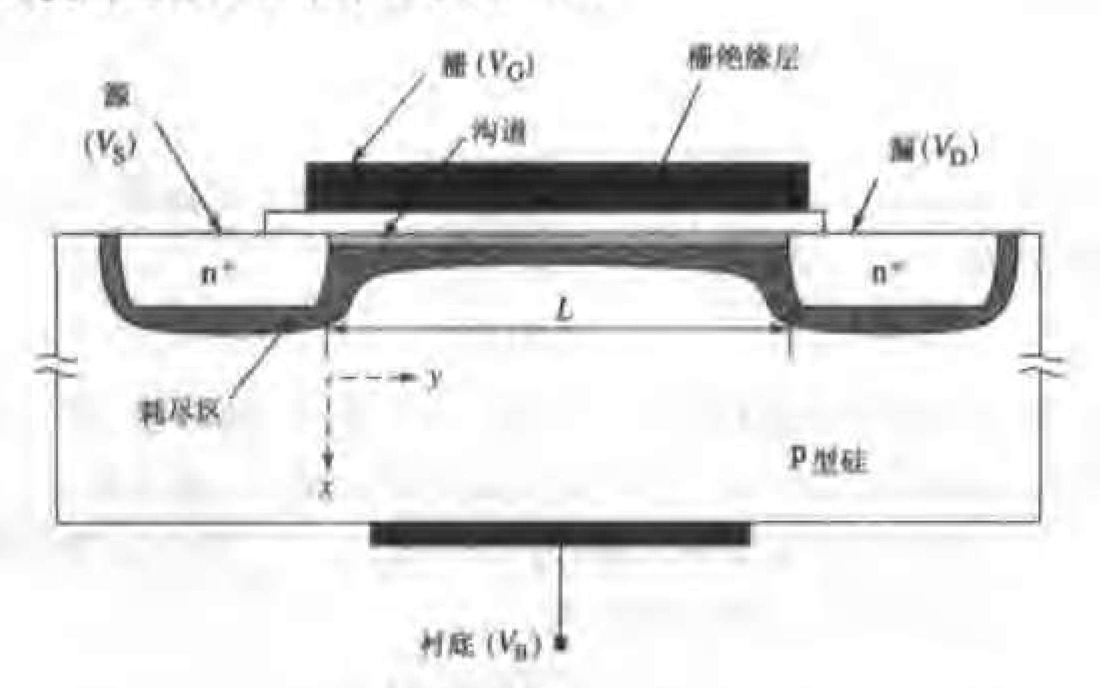


图9.1 n沟道MOSFET的基本结构。源极到漏极间的距离为沟道长度L,沟道宽度W的方向是垂直于纸面的z轴方向

MOSFET为四端元件,四个端电压的大小都对MOSFET的电学特性有影响:V_G——栅极电

压;V_S——源极电压;V_D——漏极电压;V_B——衬底电压(又称体电压)。图9.2列出了MOSFET的电学符号,其中(a)、(b)为p沟道增强型和p沟道耗尽型MOSFET,(c)、(d)为n沟道增强型和n沟道耗尽型MOSFET,从源极到衬底的二极管方向(图中箭头)指出了晶体管类型,沟道如果出现加粗的线条表示该管为耗尽管。

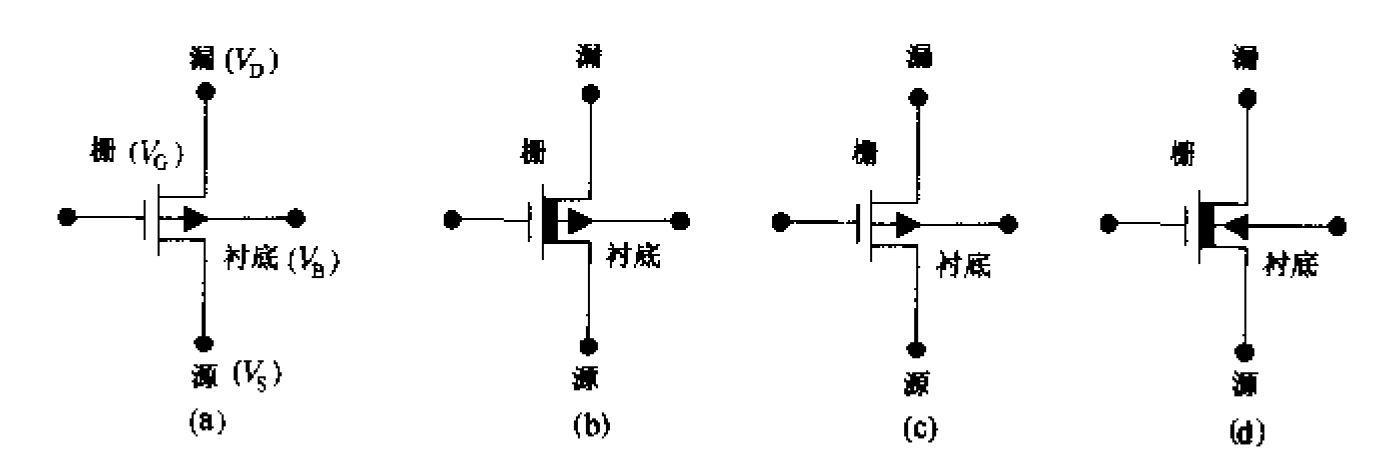


图9.2　MOSFET的电学符号:(a) p沟道增强型;(b) p沟道耗尽型;(c) n沟道增强型;(d) n沟道耗尽型

为了推导最基本的MOSFET理论,我们首先分析源极和衬底短路的情况($V_{SB}=0$)。下一节将考虑衬源偏压对MOSFET的影响,分析$V_{SB}\neq 0$的情况。

在实际应用中,MOSFET不仅可以在模拟电路用做电压控制的电阻,还可以在数字电路中用做开关。由第8章知,当n沟道MOSFET的栅极电压低于阈值电压V_T时,源漏之间的半导体表面处于耗尽或积累(空穴积累)状态。无论是哪种状态,n型源区和漏区之间都不存在导电沟道,MOSFET处于“关态”。当栅极电压增高并超过阈值电压时,栅下半导体表面出现反型区,形成包含可动电子的沟道,在源极和漏极间形成电流。继续增加栅极电压,沟道中的电子密度增大,源漏之间的电阻下降,MOSFET工作在“开态”。MOSFET从关态过渡到开态时对应的电压通常被称做阈值电压。上述分析建立在第8章的MOS电容理论上:栅极电压可以控制硅反型层中的载流子浓度。

9.1.1　强反型区

MOSFET处于强反型状态(开态)时,沟道载流子——电子的运动主要是漂移运动,流入器件的电流绝对值的大小可以表示成

$$I_D = WQ_n(y)v(y) \tag{9.1.1}$$

式中$-Q_n$代表单位面积沟道中的反型载流子电荷总量,它是y的函数。$v(y)$代表载流子的漂移速度,它也是位置y的函数。I_D和$v(y)$的正方向规定为沿$+y$轴方向。当y方向的电场强度不大时(V_D比较小),漂移速度可以表示成$v(y)=-\mu_n\mathscr{E}(y)$,其中$\mu_n$为沟道载流子迁移率,$\mathscr{E}(y)=-\partial V(y)/\partial y$,等式(9.1.1)可以表示成

$$I_D = WQ_n(y)\mu_n\partial V(y)/\partial y \tag{9.1.2}$$

如果假设:(1)反型层内载流子迁移率μ_n为常数;(2)缓变沟道近似成立,即沟道反型层和体电荷都只由垂直方向的电场强度来决定,这样第8章介绍的一维理论就可以用在x(垂直表面)方向上计算反型层的载流子浓度;(3)V_T不是y的函数。那么,$Q_n(y)$可以表示成

$$Q_n(y) = -C_{ox}[V_G - V_T - V(y)] \tag{9.1.3}$$

将等式(9.1.3)代入等式(9.1.2),沿沟道长度从源极($y=0, V_S=0$)到漏极($y=L, V_S=V_D$)积分,得到

$$\int_0^L I_D \mathrm{d}y = \mu_n W C_{ox} \int_0^{V_D} [V_G - V_T - V(y)] \mathrm{d}V \tag{9.1.4}$$

从等式(9.1.4)求解漏极电流,得到长沟 MOSFET 公式

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[\left(V_G - V_T - \frac{1}{2}V_D\right)V_D\right] \tag{9.1.5}$$

(MOSFET 的一些比较重要的公式总结在本章末的表 9.4 中) 考虑等式(9.1.5),当 $V_G > V_T$并保持 V_G不变,V_D从零不断增大时,漏极电流随 V_D线性增加;当 $V_D/2$ 大于($V_G - V_T$)时,漏极电流增长的斜率下降,如图 9.3 所示。等式(9.1.5)指出,在足够高的 V_D下,漏极电流斜率逐渐下降至零,并最终变为负数,出现负阻现象。产生这种不符合实际情况的原因是:等式(9.1.5)是由等式(9.1.3)推导出的,而等式(9.1.3)只有在 $Q_n(y)$为负数的时候才有意义,隐含的条件是 $V(y) < (V_G - V_T)$。显然 $V(y)$的最大值是 V_D,因此等式(9.1.5)只有在 $V_D \leqslant (V_G - V_T)$时候才有意义。当 $V(y) = (V_G - V_T)$时,$Q_n(y)$在 $y=L$ 处为零,这种现象称为夹断。图 9.4 为不同的漏极电压下,反型沟道和耗尽区的横截面示意图,其中包括了夹断情况。当漏极电压超过夹断电压时(图9.4(c)),导电沟道和漏区分离,沟道部分的电压降仍保持在($V_G - V_T$),夹断区(漏端到 L')的电压降为 $V_D - (V_G - V_T)$,L'称为夹断点,在该点反型载流子浓度降为零。$V_D > V_{Dsat}$时MOSFET工作在饱和区,其中 $V_{Dsat} = V_G - V_T$,并且漏极电流不再随漏极电压的升高而升高。饱和区漏极电流的大小可以通过将 $V_D = V_G - V_T$代入等式(9.1.5)得出

$$I_{Dsat} = \mu_n C_{ox} \frac{W}{2L}(V_G - V_T)^2 \tag{9.1.6}$$

由等式(9.1.5)和(9.1.6)得出的 $I_D \sim V_D$曲线如图 9.3 所示,电流 I_D随 V_D增大,直至 $V_D = V_{Dsat}$和 $I_D = I_{Dsat}$,此时夹断点坐标为 $L' = L$;继续增大漏极电压,夹断点向源区移动,漏极电流保持不变,其大小由等式(9.1.6)决定。这就是"经典"的 $I_D \sim V_D$关系,它可以比较准确地描述长沟($\geqslant 10\mu$m)MOSFET 的特性,但对于短沟器件($\leqslant 1\mu$m),该等式需要较大的修正。

现在,我们来考虑前面推导出的 MOSFET 模型的物理意义。等式(9.1.5)可以改写成

$$I_D = W C_{ox}\left(V_G - V_T - \frac{1}{2}V_D\right)\mu_n \frac{V_D}{L} \tag{9.1.7}$$

考虑等式(9.1.7),如果电压沿沟道线性下降,那么沿沟道方向电场强度为 V_D/L,可以将 $C_{ox}(V_G - V_T - V_D/2)$看做是沟道中间位置($y = L/2$)处的单位面积反型层载流子电荷总量,这样等式(9.1.5)和(9.1.7)就可以理解成沟道中的"平均"电荷浓度在恒场下的漂移流。在夹断区,$Q_n = 0$,从模型上看,电流的大小应取决于载流子通过夹断点的速度,这就产生了矛盾:沿导电沟道方向($0 \leqslant y \leqslant L'$)的电荷浓度是不断下降的(等式(9.1.3)),但又得保持电流连续,那么要求载流子的漂移速度 $v(y)$沿沟道不断提高以保证电流的大小不变。夹断点

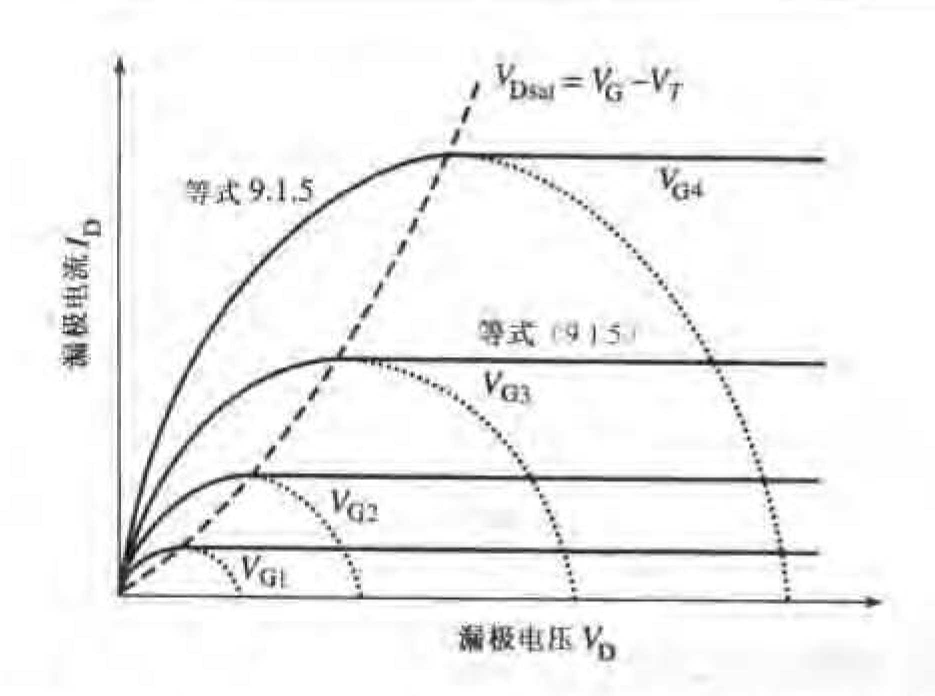


图 9.3　不同栅压下的 I_D-V_D 关系曲线。图中的点线是由等式(9.1.5)计算出的 $V_D>V_{Dsat}$ 区域。由等式(9.1.3)计算出的该区域电子浓度会小于零，不符合实际情况，因此该区域的正确表达式为等式(9.1.6)。图中栅极电压 V_G 自下向上逐渐增大，V_B 和 V_S 接零电位

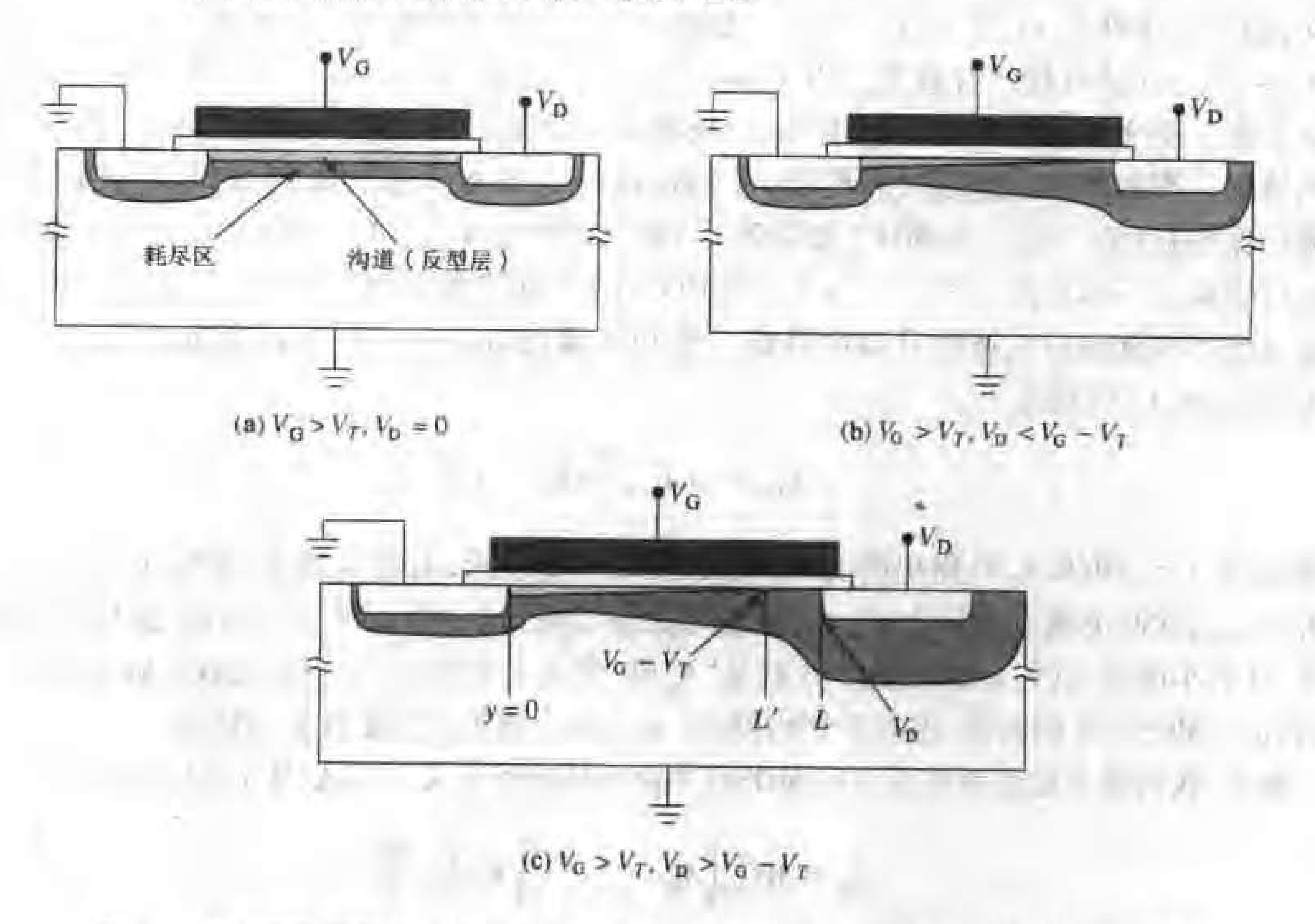


图 9.4　MOSFET 不同偏置下的横截面示意图：(a)漏极电压很小，耗尽区厚度沿沟道长度方向基本不变；(b)漏极电压增大，耗尽区厚度明显改变；(c)漏极电压超过 V_{Dsat}，沟道在 L' 处夹断，沟道长度 $L'<L$，L' 处的电位为 $V_{Dsat}=V_G-V_T$。

电荷浓度 Q_n 下降至 0，这就意味着夹断点的载流子速度必须达到无穷大，这是不可能的。对于长沟器件，这一矛盾并不明显，原因是漏极电流决定于源区和夹断点之间的沟道载流子输运速度，受夹断点和夹断区的影响不是很大。为了分析理想的 MOSFET 特性，下面给出一个形象化的物理模型。

水库模型 MOSFET的工作原理，可以用水库模型来作类似的分析。如图9.5所示，可动载流子可以类似成水分子，源区和漏区就是一个巨大的水库，它们的高度差就是源极到漏极的电压降，沟道区就像是连接两个水库的沟渠，其深度由栅-沟道电压决定。

如果源漏等电位，源漏以及沟渠中的水位就是一样的（见图9.5(a)）。当加上漏极到源极的电压时，相当于漏极的水位下降了，导致水从源区流向漏区。水流的速度随着高度差（类比于V_D）的上升而增大。我们知道，在任何一个截面的水流都必须连续，那么水流的速度就会因为沟渠深度的减小而加快。起初，通过沟渠的水流是由沟渠的深度（栅极电压控制）和两个水库之间的高度差决定的（见图9.5(b)）。沟道夹断情形可以这样类似：由于沟渠容量有限，水流大小达到了沟渠输运能力的最大值。漏区水库的水位继续下降，当漏区的水位低于沟渠的最低水位时，水流将像瀑布一样冲入漏区水库（见图9.5(c)）。水流的速率等于瀑布最高处的速率，而与瀑布的高度差（类比于$V_D - V_{Dsat}$）无关。我们可以类似地分析MOSFET，一阶模型可暂时不考虑载流子离开沟道之后，在强电场的空间电荷区中的输运机制。

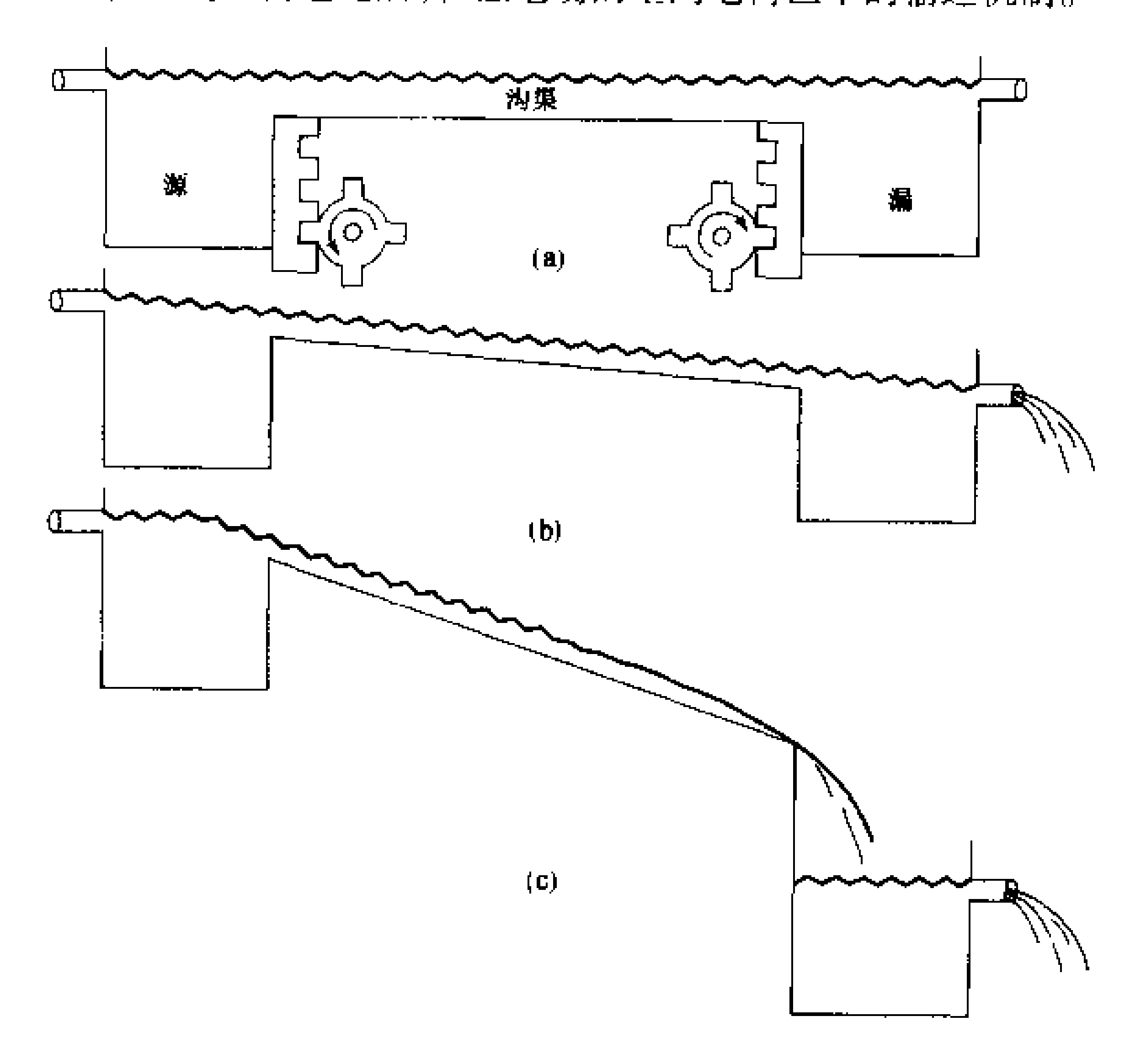


图9.5 MOSFET的水库模型。(a)源区和漏区在同一水平面上时（$V_D = 0$），没有水流，齿轮决定了沟渠中的水深（V_{GS}）；(b)漏区的水位比源区低时，水沿着沟渠从源区流向漏区；(c)沟渠的输运能力有限，水流达到最大值。继续降低漏区水位，只是增加了漏端处瀑布的高度差，水流大小不再增加

9.1.2 沟道长度调制

等式(9.1.5)被称为长沟MOSFET理论。对于长沟器件，特别是漏源电压不是很高时，可以假设沟道中的载流子迁移率为常数时，此时该等式足够精确。

迄今为止,我们只是考虑了 MOSFET 的直流情况。MOSFET 的交流(和开关)特性与沟道长度 L 有明显的关系,缩短 L 可直接改善这些特性。另外,由等式(9.1.5)知,漏源偏压不变时,缩短沟道长度可以提高漏极电流。因此,不管是早期还是现在,缩短沟道长度都是 MOSFET 设计的一个目标。当然,随着 L 的减小,长沟 MOSFET 理论需要做相应的修正以满足精度要求。短沟 MOSFET 饱和区漏极电流不再保持为常数(等式(9.1.6)中的 I_{Dsat}),而是随漏极电压的升高而升高,产生这种现象的主要机制是沟道长度调制。

沟道长度调制可以直接通过修正长沟器件的模型公式来实现。等式(9.1.6)中的漏极电流与沟道长度成反比,并且当漏源电压超过 V_{Dsat} 时,沟道长度随着电压的升高而减小。如果用 $\Delta L = L - L'$ 表示沟道长度的减小量,则

$$I_D(V_D > V_{Dsat}) = I_{Dsat}\left[\frac{L}{L - \Delta L}\right] \approx I_{Dsat}\left[1 + \frac{\Delta L}{L}\right] \tag{9.1.8}$$

式中的 ΔL 表示图 9.4(c)中 L 到 L' 的距离。随着($V_D - V_{Dsat}$)的升高,ΔL 将增大。等式(9.1.8)中的近似是考虑到 $\Delta L \ll L$,通过台劳展开取前两项得到的。通常情况下这个近似条件是满足的。然而,电路设计者需要的是电流电压关系式,因此,我们来推导 ΔL 与电压的关系,方法是解 Poisson 方程,并假设夹断区电荷密度为 N_a

$$\Delta L = \left[\frac{2\epsilon_s}{qN_a}(V_D - V_{Dsat})\right]^{1/2} \tag{9.1.9}$$

这种方法不是很准确(因为忽略了夹断区的自由电荷,并且只考虑了一维情形),而且很麻烦。实用的体现 MOSFET 沟道长度调制效应的公式采用了类似双极晶体管 Early 电压的概念,将电流表达式写成

$$I_{Dsat} = \frac{\mu W C_{ox}}{2L}(V_G - V_T)^2\left(1 + \frac{V_D}{V_A}\right) \tag{9.1.10}$$

式中 V_A 类似双极晶体管的 Early 电压(7.1 节)。

9.1.3 衬偏效应

到现在我们只考虑了衬偏为零的情况,即源极和衬底极(衬底引出端)相连,$V_S = V_B$。只要源极和衬底极之间的 pn 结零偏或反偏,通过的电流就非常小,但是反偏电压 V_{SB} 却能够影响源漏之间的电流 I_D,这种效应称为体偏置效应或衬偏效应。

我们可以利用图 9.6 中的二维图形来定性地理解这种效应。图 9.6 是两种衬偏($V_{SB} = 0$ 和 $V_{SB} = -1V$)和三种栅压(对应于平带,$2|\phi_p|$ 和反型)下的能带图。比较图 9.6(d)和图 9.6(f)知,加上衬偏电压后,需要能带更大的弯曲才能使表面反型,因此需要更大的栅极电压,也就是说阈值电压上升了。

第 8 章已经分析了衬偏电压对阈值电压的影响,并得出了阈值电压的表达式(8.3.18)。相对于零衬偏,阈值电压的改变量为

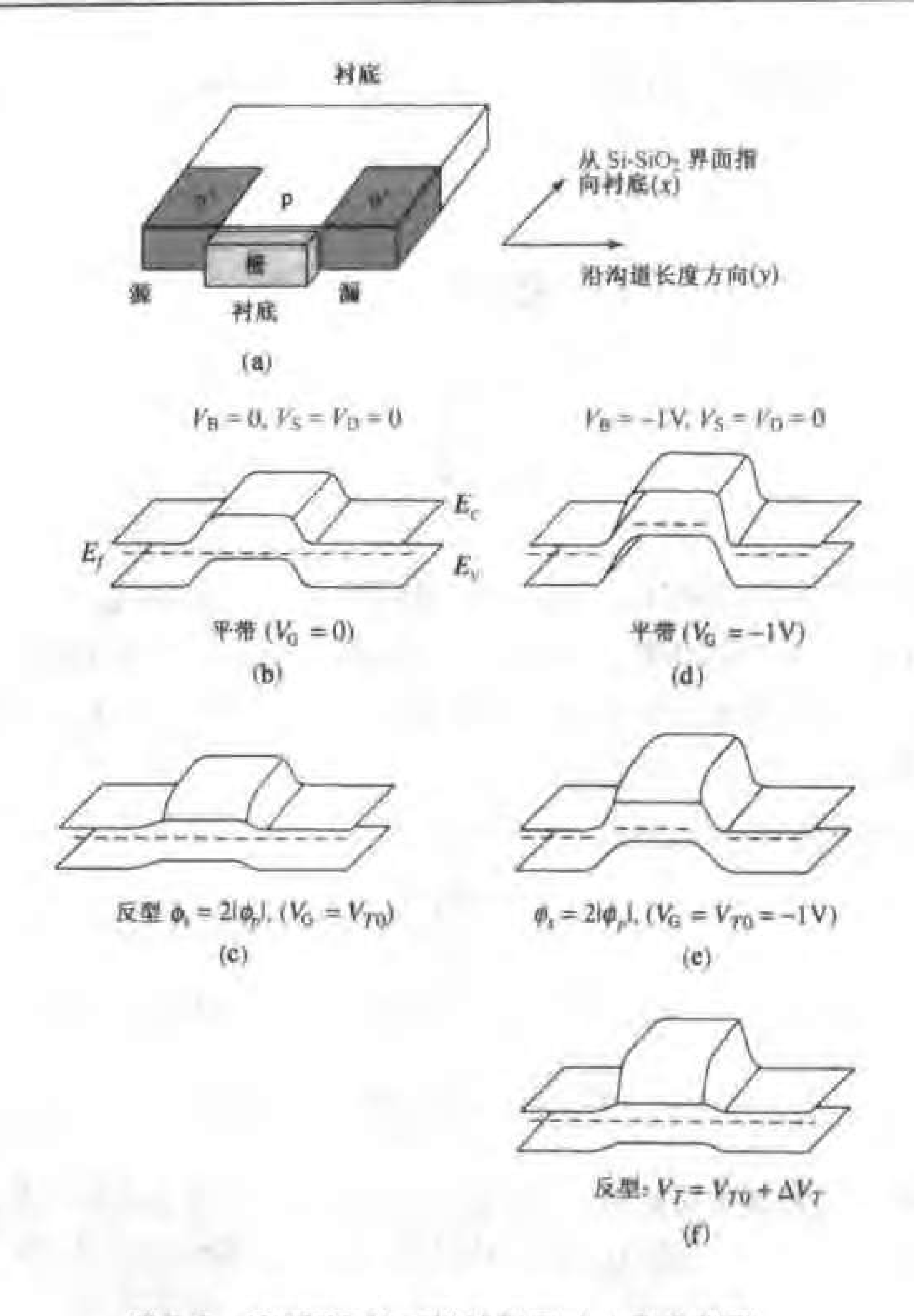


图 9.6 MOSFET 的三维结构图 (a) 和能带图,
(b)、(c): $V_B = 0V$, (d)、(e)、(f): $V_B = -1V$

$$\Delta V_T = \frac{\sqrt{2\epsilon_s qN_a}}{C_{ox}}\left(\sqrt{2|\phi_p| + |V_{SB}|} - \sqrt{2|\phi_p|}\right) = \gamma\left(\sqrt{2|\phi_p| + |V_{SB}|} - \sqrt{2|\phi_p|}\right) \tag{9.1.11}$$

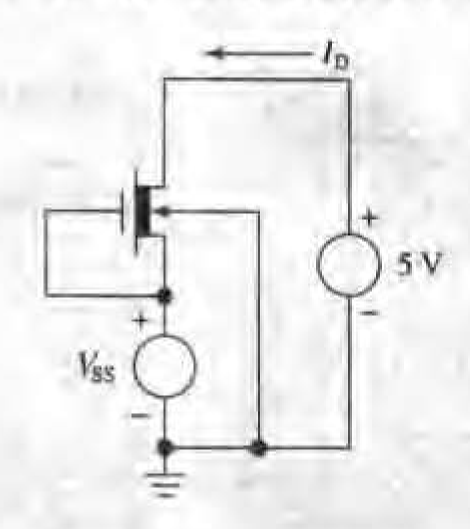
式中 γ 为体效应系数,单位是 $V^{1/2}$

$$\gamma = \frac{\sqrt{2\epsilon_s qN_a}}{C_{ox}} \tag{9.1.12}$$

例题 耗尽型 MOSFET 参数

假设等式(9.1.5)和(9.1.6)适用于阈值电压小于零的耗尽型 n 沟道 MOSFET,已知衬底掺杂浓度为 $1.63 \times 10^{15} cm^{-3}$,体效应系数 $\gamma = 0.5V^{1/2}$。连接 MOSFET 的电路如下图所示,$V_{SS} = 0V$ 时测得电流 I_D 为 30μA;当 V_{SS} 上升至 1V 的时候,I_D 减小至 23.1μA。

计算 $V_S=0V$ 时的阈值电压 $V_T(0)$,并求出等式(9.1.5)和(9.1.6)中的系数 $\mu_n C_{ox} W/L$。



解:由于上述两种情况都有 $V_{DS} \geq 4V$, $V_{GS}=0V$,我们可以假设 $V_{DS}>(V_{GS}-V_T)$,这样两种情况都满足等式(9.1.6)的条件。我们可以求出阈值电压 V_T后再来验证假设是否成立。由于源极电压改变,我们需要用等式(9.1.11)来考虑衬偏效应,因此首先要计算出 $|\phi_p|$ 值。利用衬底掺杂浓度和等式(4.2.9(b)),求出 $|\phi_p|=0.3V$。

$V_{SS}=0V$ 时,源极电位为零,并且

$$30=\mu_n C_{ox}\frac{W}{2L}[0-V_T(0)]^2$$

$V_{SS}=1V$ 时,源极电位等于 1V,用等式(9.1.11)计算出阈值电压的改变量

$$\Delta V_T=0.5\left[\sqrt{0.6+1}-\sqrt{0.6}\right]=0.245\text{ V}$$

对于 n 沟道 MOSFET,衬偏效应使得阈值电压上升(因为零偏阈值电压为负,所以绝对值下降),此时的阈值电压 $V_T(1)=V_T(0)+0.245$,代入 $V_{SS}=1V$ 的电流公式,有

$$23.1=\mu_n C_{ox}\frac{W}{2L}\left[0-(V_T(0)+0.245)\right]^2$$

因此

$$\left[\frac{I_D(0)}{I_D(1)}\right]^{1/2}=1.14=\frac{-V_T(0)}{-V_T(0)-0.245}$$

从上式可以求出 $V_T(0)=-2V$。代入任意一种情况下的电流电压关系式,可以计算出 $\mu_n C_{ox} W/L=15\mu A/V^2$。我们得出的阈值电压使得 MOSFET 在两种情况下都工作在饱和状态,$V_{DS}>(0-V_T)$,所以假设成立。

实际上,等式(9.1.11)成立的条件是 MOSFET 的沟道长度远大于源-衬反偏 pn 结的耗尽区的宽度。但对于短沟器件,解一维 Poisson 方程(得出等式(8.3.18))已不能满足要求,需要用二维理论来描述空间电荷的分布,得到更为精确的阈值电压的表达式。

9.1.4　体电荷效应

推导 MOSFET 漏极电流表达式 9.1.5 和 9.1.6 的基本假定是,表面耗尽区电离施主提供的电荷 Q_d 沿沟道方向不变。但实际情况是从源到漏随着沟道电势的不断升高,耗尽区的宽度也随之增长,因而耗尽区的电荷(也称为体电荷 Q_B)也沿着沟道方向增长,如图 9.4 所示。$Q_d(y)$ 的表达式如下

$$Q_d(y) = qN_a x_d = \sqrt{2\epsilon_s qN_a[2|\phi_p| + V(y) - V_B]} \tag{9.1.13}$$

体电荷沿沟道方向的改变意味着阈值电压以及 Q_n 都与 y 相关。考虑这些因素,代入等式(9.1.3)得到

$$Q_n(y) = C_{ox}[V_G - V_{FB} - 2|\phi_p| - V(y)] + \sqrt{2\epsilon_s qN_a[2|\phi_p| + V(y) - V_B]} \tag{9.1.14}$$

用等式(9.1.14)替代等式(9.1.3),重新考虑电流输运表达式(9.1.2),沿沟道从源($y = 0$)到漏($y = L$)积分,得到长沟 MOSFET 考虑了体电荷变化的漏极电流表达式

$$\begin{aligned} I_D = \mu_n \frac{W}{L}\Bigg\{ & C_{ox}\left(V_G - V_{FB} - 2|\phi_p| - \frac{1}{2}V_{DS}\right)V_{DS} \\ & - \frac{2}{3}\sqrt{2\epsilon_s qN_a}\left[(2|\phi_p| + V_D - V_B)^{3/2} - (2|\phi_p| + V_S - V_B)^{3/2}\right]\Bigg\} \end{aligned} \tag{9.1.15}$$

$V_S = V_B = 0$ 时,等式(9.1.15)简化成

$$\begin{aligned} I_D = \mu_n \frac{W}{L}\Bigg\{ & C_{ox}\left(V_G - V_{FB} - 2|\phi_p| - \frac{1}{2}V_{DS}\right)V_{DS} \\ & - \frac{2}{3}\sqrt{2\epsilon_s qN_a}\left[(2|\phi_p| + V_D)^{3/2} - (2|\phi_p|)^{3/2}\right]\Bigg\} \end{aligned} \tag{9.1.16}$$

等式(9.1.15)含有 3/2 次幂,使用起来很不方便。为了简化模型,引入一个拟和参数 α(称为体电荷因子),等式(9.1.15)简化成

$$I_D = WC_{ox}\left(V_G - V_T - \frac{\alpha}{2}V_{DS}\right)\mu_n \frac{V_{DS}}{L} \tag{9.1.17}$$

通常体电荷因子 α 比 1 大,典型值为 1.5。考虑体电荷效应得到的漏极电流比不考虑体电荷效应得到的要小。将等式(9.1.17)对 V_{DS} 求导并令其为零,可以得到 V_{Dsat} 的表达式

$$V_{Dsat} = \frac{V_G - V_T}{\alpha} \tag{9.1.18}$$

$V_D = V_{Dsat}$ 时 MOSFET 进入饱和区,将等式(9.1.18)代入等式(9.1.6),得到饱和区电流表达式为

$$I_{Dsat} = \frac{\mu_n}{2}\frac{W}{\alpha L}C_{ox}(V_G - V_T)^2 \tag{9.1.19}$$

9.1.5 离子注入调整阈值电压

大部分 CMOS 集成电路要求精确地控制阈值电压 V_T。由第 8 章知，阈值电压的控制可以通过调整半导体表面附近的掺杂浓度来实现。由于无法精确控制阈值电压，MOS IC，尤其是 CMOS IC 的发展滞后了很多年。直到 20 世纪 80 年代，出现了离子注入工艺，才能够比较精确地调整阈值电压的大小。利用离子注入技术(见第 2 章)，在半导体表面的指定位置可以得到所需要的掺杂浓度。离子注入是用强场加速的离化杂质原子束(能量大约在 100keV)轰击硅片表面，这些离子进入硅(或者是氧化薄层或阻挡薄层)，到达一定的深度，一般不到一个微米。停留在氧化层中的离子通常是没有活性的。注入到硅中的离子在适当的温度下退火，将成为激活的施主或者受主杂质。合理控制退火的温度和时间，前面通过扩散工艺获得的 pn 结杂质分布就不会有较大的改变。

注入可以增加或减小(通过补偿)硅表面附近的净掺杂浓度。在 n 沟道的 MOS 工艺中，需要在场区(芯片上有源器件以外的区域)注入大剂量的硼以提高场区的阈值电压。沟道区则需要轻剂量注入，用来调整阈值电压或抑制次表面穿通(参考 9.2 节)。

图 9.7 为典型的离子注入形成的掺杂原子沿深度方向的分布，该分布为 Gauss 分布(图中实线)，注入浓度 $N_i(x)$ 为

$$N_i(x) = \frac{N'}{\sqrt{2\pi}\Delta R_p}\exp\left[-\frac{(x-R_p)^2}{2(\Delta R_p)^2}\right] \tag{9.1.20}$$

其中 N' 为单位面积注入的离子数(剂量)，R_p 为注入离子的平均投影距离(射程)，标准偏差 ΔR_p 描述了分布"宽度"。射程和宽度都与注入离子束能量以及注入离子的种类有关。图 2.16、2.17 和 2.18 给出了常用的硅掺杂离子的 R_p 和 ΔR_p 的数值。

通过退火激活注入的离子后，杂质分布变缓，如图 9.7 中的虚线所示。计算离子注入对阈值电压的影响可以采用比较简单的假设，把实际的分布近似成两段分布(方盒近似)，如图 9.7 中的点线。假设注入浓度从表面到深度 x_i 处为常数 N_{ai}，那么注入剂量 $N' = N_{ai}x_i$。设衬底的掺杂浓度为 N_a，那么靠近表面处的掺杂浓度为 $(N_{ai}+N_a)$。$x > x_i$ 区域的有效掺杂浓度为 N_a。

计算离子注入后的 V_T，必须考虑两种不同的情形。第一种是注入的有效深度大于反型层的厚度。这种情形可以直接利用等式(8.3.18)计算阈值电压，只需要用 $(N_{ai}+N_a)$ 替代 N_a 即可。这种分布存在一些缺点：提高了耗尽区靠近衬底一侧的掺杂浓度，增大了衬底电容；降低了沟道区击穿电压；阈值电压受衬偏的影响更加严重(更糟糕的衬偏效应)。

第二种是注入的离子都分布在表面空间电荷区。这种情形计算阈值电压比较麻烦，但还是可以利用前面求阈值电压的方法。第一步是在空间电荷浓度突变处求解 Poisson 方程，然后利用耗尽近似，求出耗尽区宽度 x_d 与表面势 ϕ_s 的关系式

$$x_d = \sqrt{\frac{2\epsilon_s}{qN_a}(\phi_s + |\phi_p|) - x_i^2\frac{N_{ai}}{N_a}} \tag{9.1.21}$$

等式(9.1.21)有实根的条件是根号下的第二项小于第一项，这和我们的假设是一致的，即第二种情形的注入深度小于耗尽区宽度。为了求出阈值电压，等式(9.1.21)中的表面势必

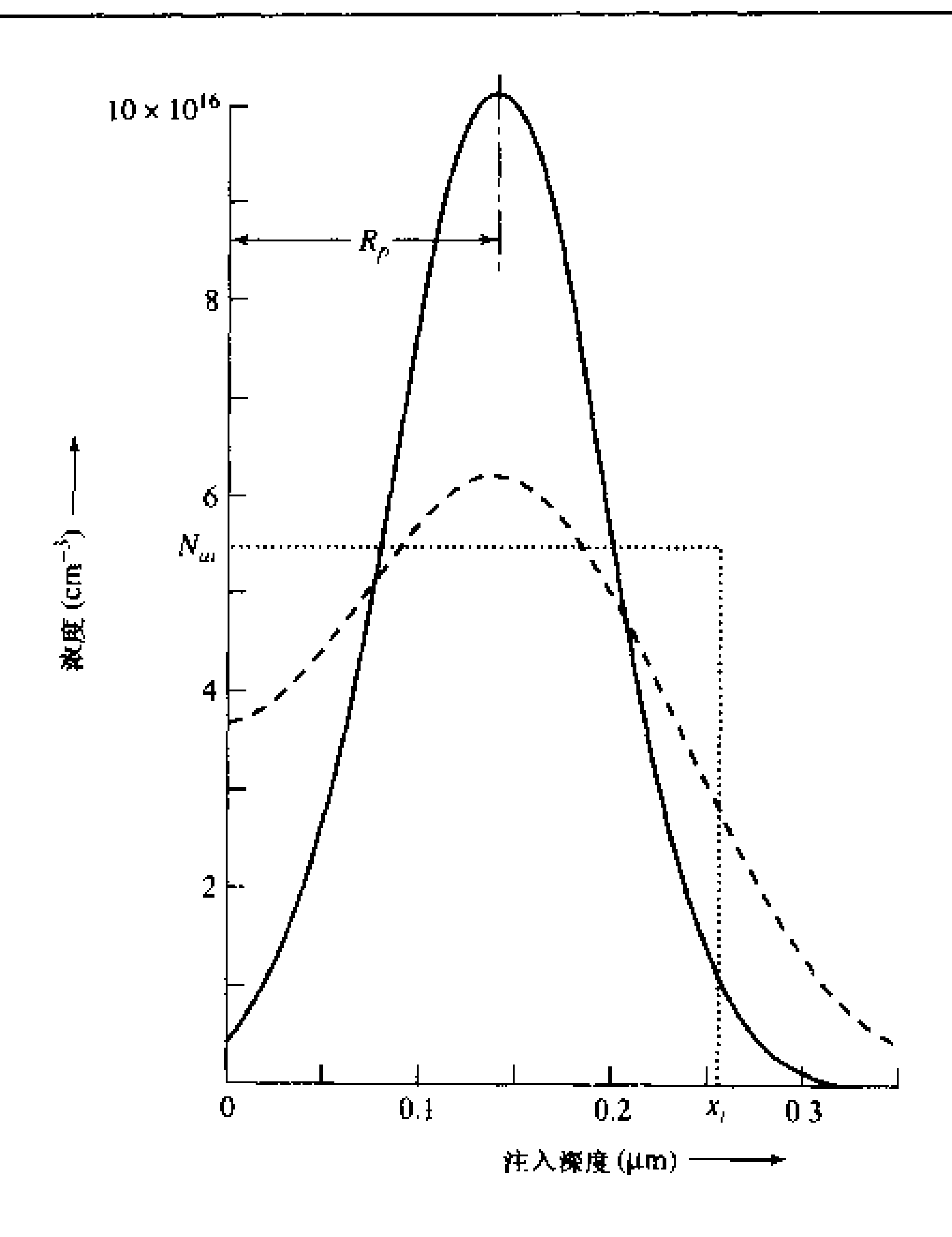


图9.7 离子注入后的杂质分布(实线)以及杂质激活后的分布(虚线) 点线分布用于计算阈值电压

须根据新的掺杂情况进行修正:$\phi_{ps}=(kT/q)\ln[(N_{ai}+N_a)/n_i]$。由等式(9.1.21)计算出$x_{d\max}$,并由此得出反型时的耗尽电荷密度$Q_d$。

$$\begin{aligned}Q_d &= -qN_{ai}x_i - qN_a x_{d\max}\\ &= -qN_{ai}x_i - \sqrt{2qN_a\epsilon_s(|\phi_{ps}|+|\phi_p|+V_{SB})-q^2x_i^2N_aN_{ai}}\end{aligned} \tag{9.1.22}$$

利用$N_{ai}=N'/x_i$,很容易写出耗尽电荷密度Q_d与注入剂量的关系,其中N'表示注入到硅中的单位面积原子密度。由8.3节阈值电压与Q_d的关系,求出阈值电压的表达式为

$$V_T = V_{FB} + V_S + |\phi_p| + |\phi_{ps}| + \frac{qN'}{C_{ox}} + \frac{1}{C_{ox}}\sqrt{2qN_a\epsilon_s(|\phi_{ps}|+|\phi_p|+V_{SB})-q^2x_iN_aN'} \tag{9.1.23}$$

阈值电压表达式(9.1.23)与参考点的选取无关。

比较等式(9.1.23)和等式(8.3.18),我们可以发现离子注入的非均匀分布对阈值电压的影响主要体现在三个方面:(1)反型时降落在表面耗尽区的压降为($|\phi_{ps}|+|\phi_p|+V_{SB}$),而不是($2|\phi_p|+V_{SB}$);(2)阈值电压与注入剂量成线性关系,见$qN'/C_{ox}$项;(3)根号下的耗尽区电荷项增加了$-q^2x_iN_aN'$一项。在这三种变化中,第一种是对数关系,影响不大;第二种是对注入剂量N'的线性依赖关系,该影响最为重要,N'项可以看做是增加了表面固定电荷的数量,将

平带电压改变成 $V_{FB}=\Phi_{MS}-Q_f/C_{ox}$(不考虑氧化层中的电荷),从而改变了阈值电压;第三种变化是阈值电压与注入深度 x_I 的关系,从图9.8可看出它们之间的依赖关系并不强。图9.8中的 V_T 由等式(9.1.23)算出,相关的工艺参数见图9.8的说明。由于对 x_I 的依赖关系不是很强,我们可以忽略平方根中的注入项,合理估算出 V_T 的大小。实践中我们的确也是这样做的。图9.8中的MOSFET不做离子注入时的阈值电压为 −0.19V,太接近于零,不满足n沟道MOSFET的可靠性要求。可见离子注入调整阈值电压是非常重要的。

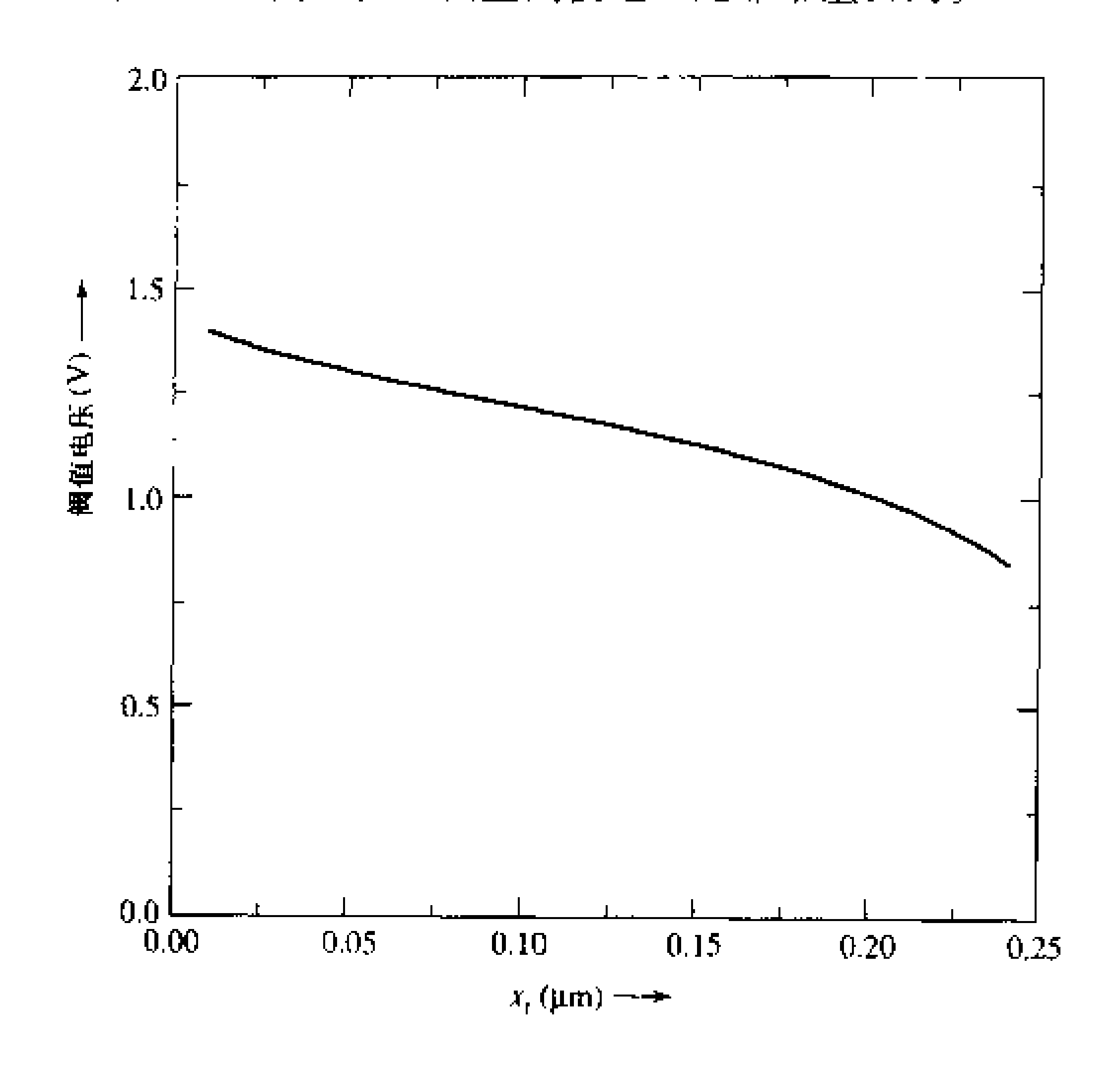


图9.8 等式(9.1.23)得出的阈值电压和注入深度 x_I 的关系,MOSFET的栅氧厚度为87nm,$Q_f/q=10^{11}\text{cm}^{-2}$,$N'=3.5\times10^{11}\text{cm}^{-2}$,$N_a=2\times10^{15}\text{cm}^{-3}$,$V_S=V_B=0$

图9.9说明了衬偏效应和离子注入对n沟道MOSFET阈值电压的调整作用。点线代表注入不同剂量的p型硼杂质,在不同的衬偏电压 V_{SB} 下测得的阈值电压。实线是以 V_S 作为参考点,由等式(9.1.23)计算得到的曲线。图9.9中的 x 轴是等式(9.1.23)中的平方根项,所以计算曲线是一条直线。衬底是(100)p型硅,硼的掺杂浓度为 $1.2\times10^{16}\text{cm}^{-3}$,氧化层厚度为100nm,氧化层固定电荷面密度 $Q_f/q=8\times10^{10}\text{cm}^{-2}$,$\Phi_{MS}=-0.92\text{V}$。从图中我们可以看出,如果提高注入剂量,则需要更大的衬偏电压使耗尽区超过离子注入的深度,这与我们的假设相符合。

9.1.6 耗尽型MOSFET

耗尽型MOSFET在 $V_{GS}=0$ 时,源漏之间就存在导电沟道。这种零偏置就存在沟道的器件常用做数字反相器的负载,因为它可以获得较高的输出电压高电平、较快的上升速度和较好的波形。由于反相器电路中的驱动晶体管必须是增强型器件,因此使用耗尽型负载要求是在同

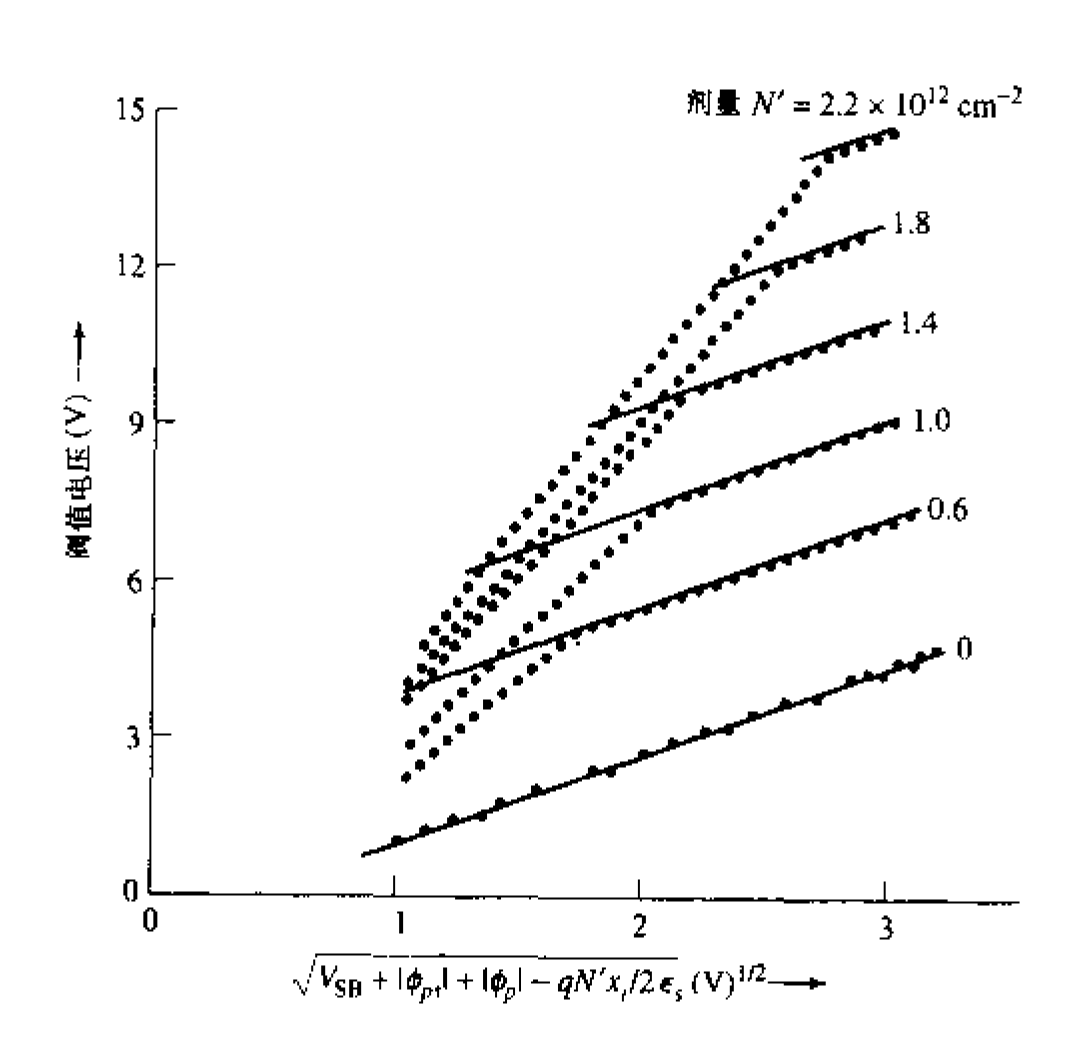


图9.9 不同的离子注入剂量的n沟道MOSFET的阈值电压(以源端为参考点)。点线为不同V_{SB}下的测量值,实线用等式(9.1.23)算出。正文中已给出了MOSFET的其他参数。(引自Hewlett-Packard Company)

一IC工艺中制造不同阈值电压的器件。选择性的离子注入方法使得我们可以制造增强型驱动/耗尽型负载(E/D)MOSFET电路。

简单来看,耗尽型晶体管就是改变了阈值电压的符号。例如,当n沟道增强型MOSFET中的施主杂质浓度不断增加时,它原本正的阈值电压数值上会减小并最终改变符号。如果只考虑阈值电压的变化,并且注入到硅中的施主杂质只位于硅-二氧化硅界面附近很薄的一层,本章前面的漏极电流模型只需改变等式中的V_T符号即可很好地适用于耗尽型MOSFET。在这种条件下,离子注入就类似于改变了表面的固定电荷浓度Q_f。对于n沟道MOSFET,离子注入后阈值电压V_{Tn}的改变可以表示成

$$\Delta V_{Tn} = -qN'/C_{ox} \tag{9.1.24}$$

对于p沟道MOSFET,注入受主杂质使阈值电压升高,改变值为

$$\Delta V_{Tp} = +qN'/C_{ox} \tag{9.1.25}$$

尽管等式(9.1.24)和(9.1.25)是近似表达式,但对于电路设计者来说,只通过阈值电压的改变量来模拟耗尽型MOSFET就足够了,尤其是用做不太关键的负载元件。

9.1.7 亚阈值区

MOSFET的一级近似认为,只有当栅极电压大于阈值电压V_T时,才能有漏极电流流过晶体管,这对大多数实际情况是适用的。但是在某些重要应用中,非常小的电流也是不能忽略

的。例如8.7节提到的,利用MOSFET的电容存储电荷的动态RAM单元,即使是非常小的电流,也会使电容放电,导致存储信息丢失。因此,我们需要重新考虑前面的的假设:V_G下降到V_T的时候沟道突然地完全关断。$V_G < V_T$时的漏极小电流称为亚阈值电流。

第8章指出,MOS表面的可动电荷密度Q_n随表面势ϕ_s指数变化,因此,当栅电压下降至低于阈值电压时,可动电荷密度并不会随之下降至零并保持为零,而是指数下降,漏极电流也是如此。利用第8章的理论,我们可以求出该区域的电流表达式。但是这样求出的表达式过于繁杂,不适合于电路设计时的计算。幸运的是,通过合理的近似,可以推导出比较直观的亚阈值电流表达式。

在亚阈值区,漏极电压几乎全部降落在反向偏置的漏衬结耗尽区上,因此,漏极电流的漂移流部分可以忽略。另一方面,自由载流子浓度沿沟道方向的梯度相当大,因此,亚阈值电流的主要成分是扩散流,类似于基区均匀掺杂的双极晶体管的基区电流。从图9.10中的能带图知,加在栅极上的电压使半导体表面能带弯曲,降低了从源区到沟道区的电子势垒$q\phi_B$,电子从重掺杂的源区(类似双极晶体管的发射区)注入到p型表面区(类似BJT的基区),大部分注入的电子被漏区收集(类似BJT的集电区)。由于势垒$q\phi_B$的作用,当栅极电压低于阈值电压V_T大约0.2V时,漏极电流开始随栅极电压V_G指数变化,亚阈值电流可以写成

$$I_D \approx I_{D0}\exp\left(\frac{qV_{GS}}{nkT}\right) \tag{9.1.26}$$

在半对数图上,亚阈值电流随V_{GS}线性变化,直至非常的低"噪声"电流区,这与等式(9.1.26)一致。尽管MOSFET在亚阈值区和双极晶体管导电的物理机制很相似,但它们之间还是存在比较大的差别。首先,MOSFET亚阈值导电的注入是局部的,只发生在表面区域;其次,BJT中基极-发射极偏压V_{BE}是外加的,MOSFET与之类似的电压是沟道靠近源区一端的表面能带弯曲ϕ_s和源极电压V_S的差值。

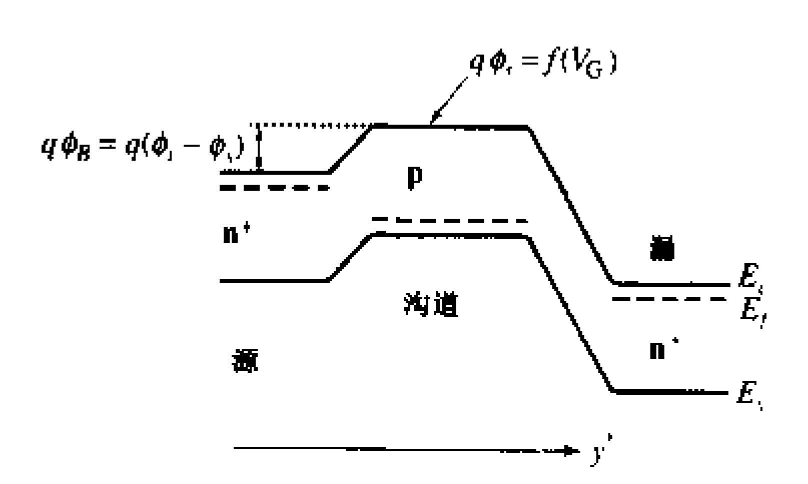


图9.10　n沟道MOSFET沟道区的能带图,源端势垒高度$q\phi_B$与栅极电压有关

我们用参数η来表示栅电压和表面势的关系

$$\eta = \frac{d\phi_s}{dV_{GS}} = \frac{1}{1 + (dV_{ox}/d\phi_s)} = \frac{1}{1 + (C_d/C_{ox})} \tag{9.1.27}$$

式中C_d是表面耗尽层电容($C_d = \epsilon_s/x_d$)。比较等式(9.1.26)和(9.1.27),我们发现"效率"因子$\eta = 1/n$,所以亚阈值电流随着η的增大而指数增大。图9.11是测量得到的亚阈值电流数

据，半对数坐标系中亚阈值电流为直线。直线斜率的倒数 S 定义为亚阈值斜率倒数（通常简称为亚阈值斜率），它代表亚阈值电流每变化 10 倍所要求的栅压变化量（单位为 mV）。根据 S 的定义，有

$$S = \frac{1}{\eta}\frac{kT}{q}\ln 10 = n\frac{kT}{q}\ln 10 \approx 60\,n\frac{\text{mV}}{\text{decade}}, \quad 25°\text{C时} \tag{9.1.28}$$

式中

$$n = 1 + \frac{C_d}{C_{ox}} \tag{9.1.29}$$

等式(9.1.29)适用于理想表面（没有表面陷阱）。如果表面存在陷阱，则等式(9.1.29)中需要加入另一项来表示栅压对表面势控制作用的降低（见 8.5 节）。

室温(25℃)下，现代典型工艺制造的 MOSFET 的 S 值的范围为 70 ~ 120mV/decade。对设计者来说，S 的大小非常重要，因为设计者需要用它来确定保证 MOSFET 处在"关"态所需要的栅压。例如，MOSFET 处在"关"态的标准是当 $V_G = 0$ 时流过漏极的电流不超过 $V_G = V_T$ 时漏极电流的 0.001%（10^{-5}）。如果给定了 S，就可以求出 V_T 的下限。

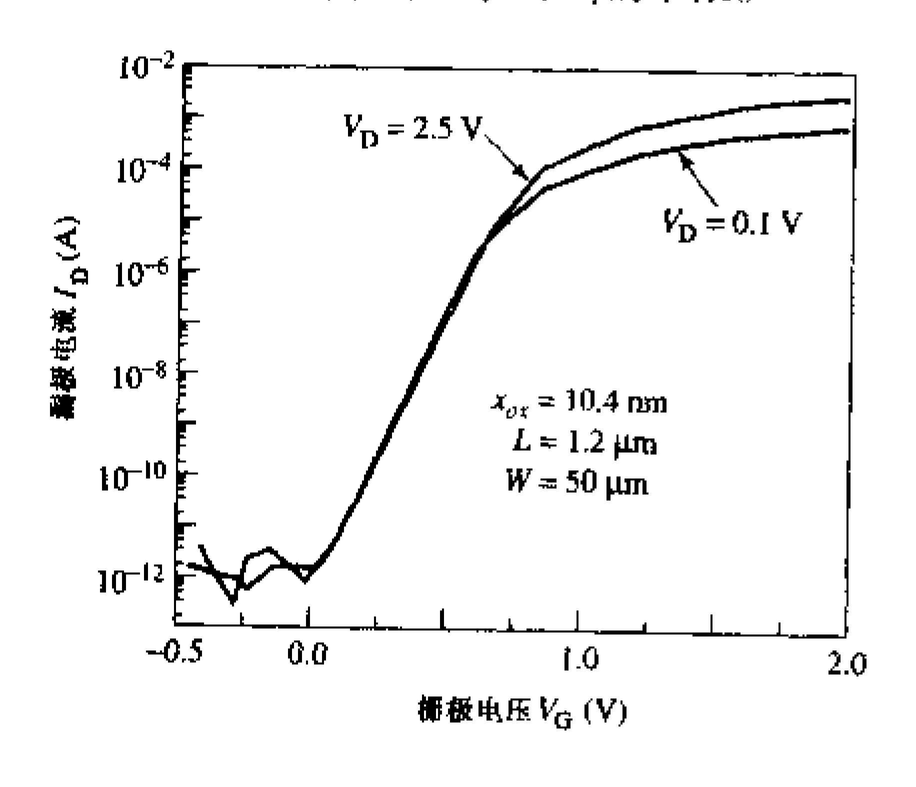


图 9.11　亚阈值电流测量曲线，MOSFET 沟道长度为 1.2μm。在半对数坐标中，图中直线部分斜率的倒数称为亚阈值斜率 S（单位为 mV/dec）

例题　亚阈值电流

已知 n 沟道 MOSFET，p 型衬底的受主掺杂浓度为 $8 \times 10^{16}\text{cm}^{-3}$，$n^+$ 多晶硅栅，氧化层厚度为 15nm，栅长为 0.8μm，请计算 $V_G = V_T$ 时的漏极电流是 $V_G = 0$ 时的漏极电流的多少倍。

解：查表 8.1，并从 8.3 节知，对于 n^+ 多晶硅栅，衬底的受主掺杂浓度为 $N_a = 8 \times 10^{16}$ cm^{-3}，那么平带电压和 ϕ_p 分别为 −0.9V 和 0.36V。

栅氧化层电容 $C_{ox} = 2.3 \times 10^{-7}\text{F/cm}^2$，利用等式(8.3.18)计算出阈值电压为 0.41V。从等式(8.3.6)计算出 $x_{dmax} = 1.28 \times 10^{-5}\text{cm} = 0.128\mu\text{m}$，因此，$C_d = 8.2 \times 10^{-8}\text{F/cm}^2$。

从等式(9.1.29)求出 $n = 1.36$，$S = 82\text{mV/dec}$

已经求出阈值电压 $V_T=0.41\text{V}$,因此,$V_G=V_T$时的漏极电流相对于 $V_G=0$ 时的漏极电流下降的数量级为(0.41/0.082)dec。也就是说,$V_G=V_T$时的漏极电流是 $V_G=0$ 时的漏极电流的 10^5倍。

在 MOS 数字电路设计中,对 ON/OFF 电流比的要求限制了可用阈值电压的范围。较低的阈值电压可以降低对电源电压的要求。为达到某些电路要求的性能,可以用更为精细的工艺制造出两种不同阈值电压的 n 沟道 MOSFET。这样,大部分晶体管的阈值电压较高,限制了关态电流,从而降低了静态功耗。但是也有较小一部分晶体管需要在开启状态下能够提供较大的电流(保证高速工作),这些器件的阈值电压就设计的比较低。为了获得高性能,必须付出的代价是较高的关态电流。为了减小这些晶体管的泄漏电流,我们可以在这些晶体管不工作的时候,加上衬偏电压来提高它们的阈值电压。

MOSFET 的速度受到两个内在因素的限制:首先,在所有的电流放大器中,最基本的是渡越时间限制(7.3 节已讨论过),即电荷沿沟道输运需要的时间;其次,由于器件本身的结构决定了存在本征电容(参见小信号模型),这些电容的充、放电需要时间。在实际应用中,还有第三个不可避免的因素(通常是最麻烦的),即寄生电容的充、放电。这些寄生电容存在于本征器件外部,模拟这些时间效应,通常需要在器件外部电路节点上加上一些电路元件。在本章的后半部,我们将继续考虑这些寄生效应。

本征 MOSFET 的速度与 MOSFET 的偏置有关。我们只考虑饱和区,因为大多数情况下,MOSFET 都工作在饱和区。首先找出沿沟道方向电场$\mathscr{E}(y)$的近似解。长沟器件的沟道电场可以直接写出。由等式(9.1.4)得

$$\int_0^y I_D \mathrm{d}y = \mu_n W C_{ox} \int_0^V [V_G - V_T - V(y)] \mathrm{d}V \tag{9.1.30}$$

对等式(9.1.30)积分,并解出 $V(y)$

$$V(y) = (V_G - V_T) - \sqrt{(V_G - V_T)^2 - \frac{2I_D y}{\mu_n W C_{ox}}} \tag{9.1.31}$$

等式(9.1.30)和等式(9.1.31)中,假设阈值电压与 y 无关,也就是说,不考虑耗尽层电荷 Q_d沿 y 方向的变化。对等式(9.1.31)求导得出电场$\mathscr{E}(y)=-\mathrm{d}V/\mathrm{d}y$

$$\mathscr{E}_y = -\frac{(V_G - V_T)}{2L} \frac{1}{\sqrt{1 - y/L}} \tag{9.1.32}$$

利用等式(9.1.32)可计算出沿沟道的渡越时间 T_{tr}

$$T_{tr} = \int_0^L \frac{1}{v_y} \mathrm{d}y = -\int_0^L \frac{1}{\mu_n \mathscr{E}_y(y)} \mathrm{d}y \tag{9.1.33}$$

并推导出

$$T_{tr} = \frac{4}{3} \frac{L^2}{\mu_n (V_G - V_T)} \tag{9.1.34}$$

尽管上述渡越时间T_{tr}的分析包含了很多近似,也没有考虑沟道中速度饱和的影响,但是计算得到的沟道渡越时间远小于实际的MOSFET电路中最快的开关时间。因此我们可以得出这样一个结论:实际MOSFET的速度与沟道渡越时间的关系不大,而主要取决于电容的充放电时间,不管是本征电容还是寄生电容。出于这个原因,在MOSFET的瞬时计算时,可以假设MOSFET的电流由器件的静态公式决定。

9.1.8 小信号等效电路模型

图9.12所示的MOSFET小信号等效电路反映了MOSFET最重要的几个效应。与栅极相关的四个电容中,只有C_{GS}和C_{GD}是本征电容,体现了栅极对沟道电荷以及电流的控制作用,是"有利"的电容,它们的大小与偏置有关。V_{DS}很小时,电容值约为$C_{ox}WL/2$;MOSFET工作在饱和区时,C_{GS}变为$(2/3)C_{ox}/WL$,而C_{GD}接近于零,表明此时栅极和漏极之间的电力线几乎为零。C_{GS}和C_{GD}充放电的速度与沟道电荷渡越时间有关。

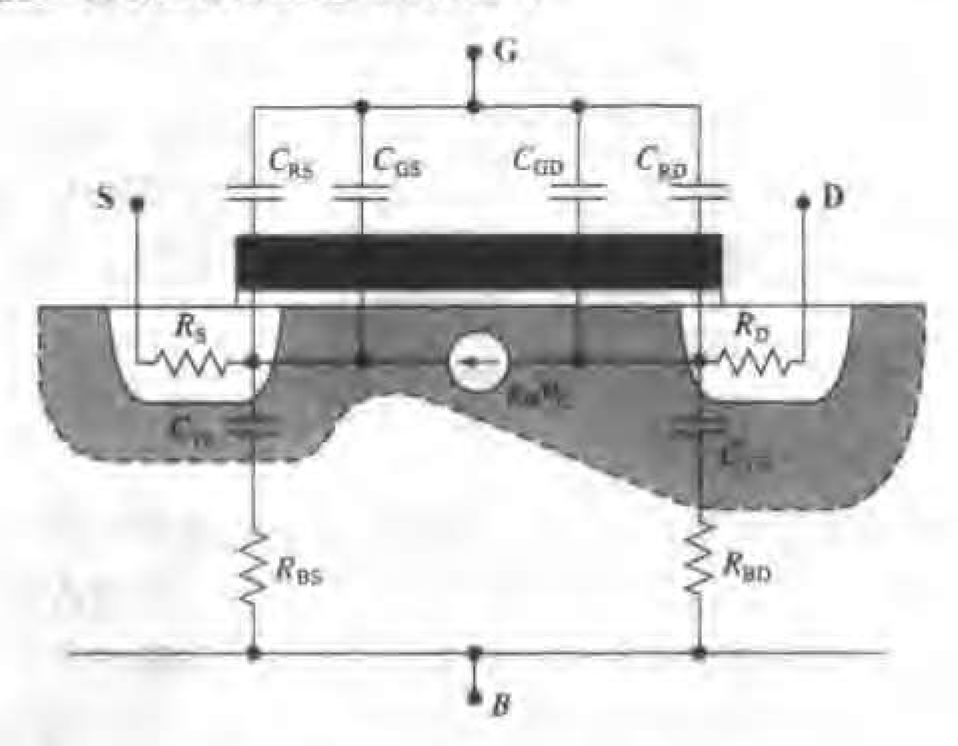


图9.12 MOSFET小信号等效电路

栅源和栅漏之间的电容C_{RS}和C_{RD}属于寄生电容,来源于栅区与源漏扩散区的重叠部分。连接衬底和源漏扩散区的电容C_{TS}和C_{TD}是反偏pn结势垒电容。电阻R_S和R_D反映了电极到沟道的串联电阻,电阻R_{BS}和R_{BD}代表源漏耗尽区边缘与衬底电极之间的欧姆电阻。

跨导g_m将晶体管的输出和输入联系起来。MOSFET的输出是漏极电流,输入是栅极电压,定义g_m为漏极电流对栅极电压的求导

$$g_m \equiv \partial I_D/\partial V_G \tag{9.1.35}$$

$V_D < V_{Dsat}$时,由等式(9.1.5)得

$$g_m = \mu_n C_{ox}(W/L)V_{DS} \tag{9.1.36}$$

在这个区域,g_m随V_{DS}的增大而线性增大,并与栅压无关。g_m还与MOSFET的宽长比、迁移率以及单位面积栅氧化层电容成正比。当V_{GS}固定时,其中任意一项增大,都会增大漏极电流。当$V_D > V_{Dsat}$时,由等式(9.1.35)和(9.1.6)求出跨导的为

$$g_{msat} = \mu_n C_{ox}(W/L)(V_G - V_T) = \frac{2I_{Dsat}}{(V_G - V_T)} \tag{9.1.37}$$

饱和区跨导 g_{msat} 与 V_D 无关,与 V_G 成线性关系。

9.2 短沟 MOSFET 的改进模型

9.2.1 长沟分析的局限性

第一个商用 MOSFET 的沟道长度大于 20μm,用长沟模型能得到非常好的结果,尽管该模型存在一些不符合实际的结论(例如在漏端出现载流子速度达到无穷大)。长沟理论认为当栅极电压大于阈值电压时,漏极电流随着漏极电压的升高而升高,因为沟道中的载流子速度会随着沟道电场的增强而增大。当漏极电压足够高时,出现漏端"夹断",漏区附近载流子浓度降为零。然而电流必须连续,因此"夹断"区就需要载流子速度无穷大,以弥补几乎为零的载流子浓度。尽管这是不可能的,但是在沟道长度大于 10μm 的时候,这样的假设所得出的结论与实际符合很好。可是,10μm 的沟道长度远大于现在制造的大多数 MOSFET。对于这些小尺寸器件,我们需要考虑其他的物理效应,才能很好地模拟这些器件。下面我们将定性讨论这些效应。

9.2.2 短沟效应

这里我们提到的短沟效应的概念稍微有些模糊,因为它涉及到的一些二阶效应在长沟器件中也存在,比如迁移率下降、速度饱和等。严格定义的短沟效应应当是只有在短沟器件中才能观察到的现象。我们通过比较沟道长度 L 与漏衬 pn 结空间电荷区厚度来区分长和短,当二者为同一数量级时,该器件称为短沟器件。本书所说的短沟效应指的是随着 L 的减小和漏极电压的上升而出现的阈值电压下降的现象。V_T 的下降是三个因素共同作用的结果:(a)源漏电荷共享;(b)漏致势垒降低(DIBL);(c)次表面穿通。下面我们将分别讨论这些效应。

源漏电荷共享 图 9.13 是一个短沟 MOSFET,偏置状态为 $V_S = V_B$,V_D 很小(接近于 V_S),$V_G = V_T$。由第 8 章知,此时衬底耗尽区电荷 $Q_d = -qN_a x_{dmax} WL$,其中 W 指的是图中 z 方向 MOSFET 的沟道宽度,L 指沟道长度。这些电荷在图 9.13 中用斜纹线表示[28]。我们注意到该矩形区域有两个部分的电荷还属于源衬和漏衬 pn 结耗尽区,因此不需要栅极电压作用在这部分区域,用来耗尽可动载流子。长沟器件中,这两个近似于三角形的区域与整个矩形耗尽区相比可以忽略。但是随着沟道长度的缩短,这一部分的作用越来越重要。由于短沟器件中的源漏耗尽区提供了一部分体电荷,需要栅极电压感应的体电荷量总量减少,因此表面反型所需要的栅极电压减小(即阈值电压减小),如图 9.14 所示。这种短沟效应在栅氧化层较厚时更为明显,因为此时栅和沟道的耦合作用减弱,栅氧化层电容与源漏 pn 结电容的比值减小。严格求解 Poisson 方程得出,当 L 足够小时,V_T 和 L 成指数关系。由于数学推导过于繁杂[14],我们不再讨论。下面的例题通过几何近似分析,得到短沟效应的半定量概念[28]。

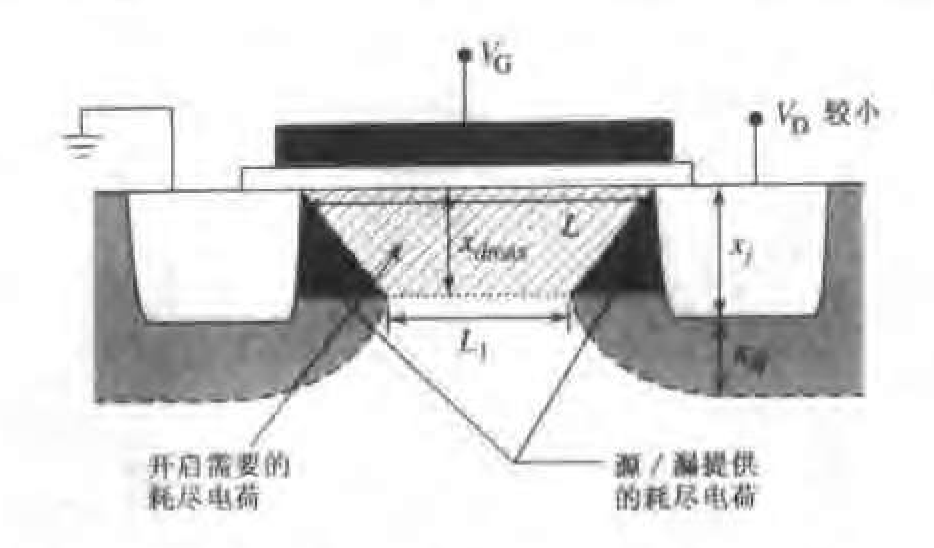


图9.13 $V_G=V_T$ 时的耗尽区电荷。矩形斜纹线区域代表达到阈值所要求的耗尽电荷 Q_d,其中部分电荷与源漏耦合,因此,栅极电压需要感应出的电荷减少,导致阈值电压降低

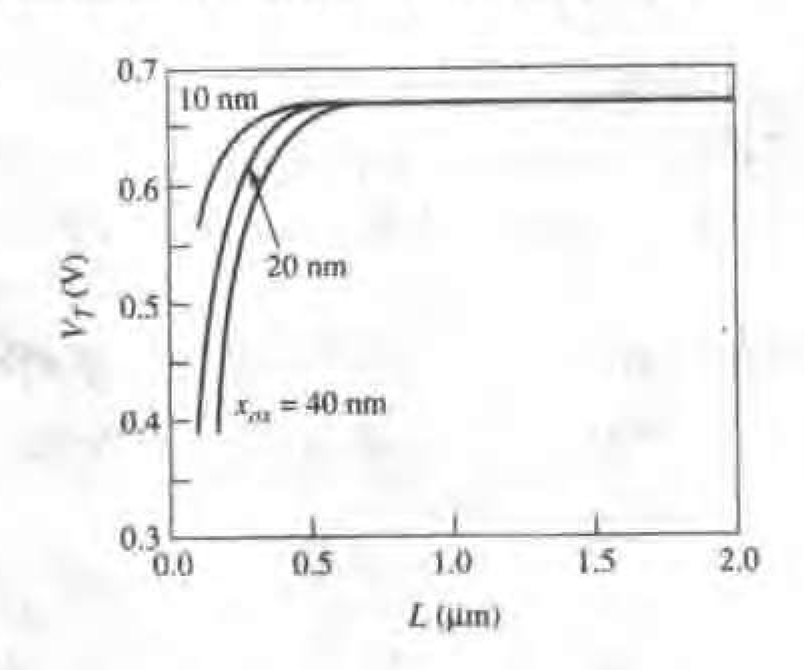


图9.14 栅氧化层厚度不同的MOSFET的阈值电压与沟道长度关系的模拟曲线

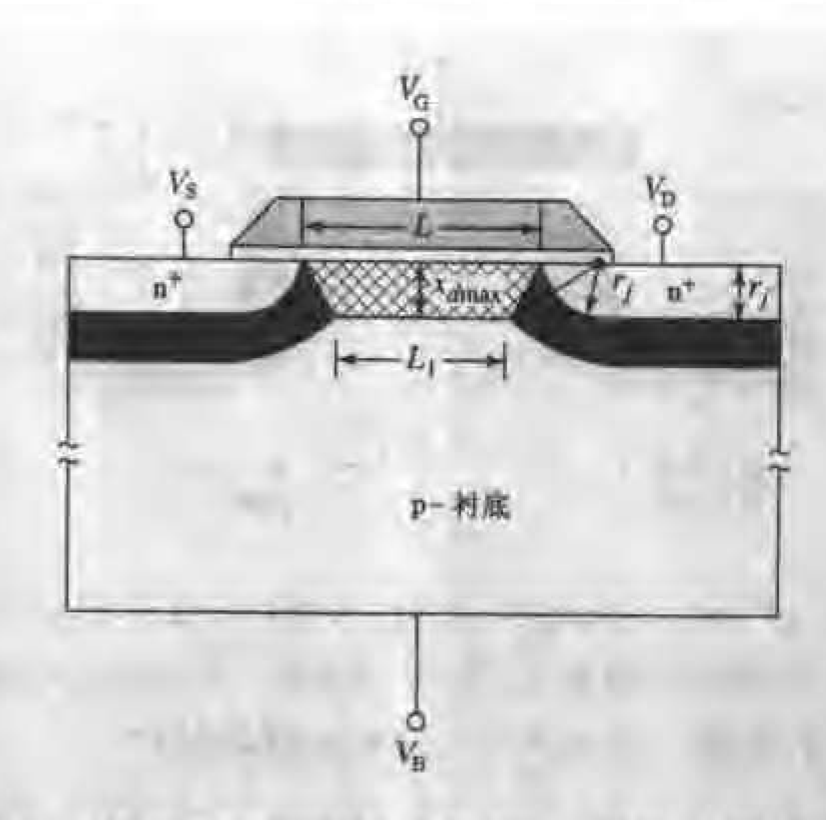
例题 电荷共享

下图是一个理想短沟MOSFET的横截面示意图,远离源漏pn结的沟道区最大耗尽层厚度为 x_{dmax},r_j是冶金结圆柱体的半径,r_2是到梯形顶点的径向距离,沟道的宽度为 W,衬底掺杂浓度为 N_a。

考虑由栅压感应的耗尽区电荷的减少,在 V_{DS}较小的条件下,推导阈值电压的近似表达式。

解:V_{DS}较小时,由 V_G引起的感应电荷横截面近似为图中所示的梯形(图中网纹区),高度为 x_{dmax},上底(表面一侧)和下底(衬底一侧)的边长分别为 L 和 L_1,那么该区域的电荷 Q_{d1} 为

$$Q_{d1}=qx_{dmax}WN_a\frac{L+L_1}{2} \tag{1}$$



等式(1)中计算出来的 Q_{d1},就是表面能够形成沟道时,栅压必须感应出的耗尽层电荷总量。如果沟道足够长,在源漏两端的空间电荷区的宽度和 L_1 相比可以忽略时,L_1 和 L 近似相等,这样,$Q_{d1}=Q_d=qx_{dmax}N_aWL$,和9.1节中的一阶模型相同。但对于短沟器件,$L_1$ 小于 L,因此 Q_{d1} 比 Q_d 小。

实用模型中,L_1 必须和实际MOSFET的几何尺寸相联系。我们可以这样近似,当 $V_G=V_T$ 时,不论是沿 x 方向(垂至于Si-SiO$_2$界面),还是沿源漏扩散区半径方向,耗尽区厚度都为 x_{dmax}。这样,$r_2=r_j+x_{dmax}$,由器件的几何结构得

$$f=\frac{Q_{d1}}{Q_d}=1-\frac{r_j}{L}\left(\sqrt{1+\frac{2x_{dmax}}{r_j}}-1\right) \tag{2}$$

参数 f 是MOSFET几何结构的函数。利用等式(8.3.18)可以直接写出阈值电压的表达式

$$\begin{aligned}V_T&=V_{FB}+2|\phi_p|+V_S-\frac{fQ_d}{C_{ox}}\\&=V_{FB}+2|\phi_p|+V_S+\frac{f}{C_{ox}}\sqrt{2\varepsilon_s qN_a(2|\phi_p|+V_S-V_B)}\end{aligned} \tag{3}$$

尽管上述推导有近似的成分,但等式3还是可以比较准确地预测短沟器件阈值电压 V_T 的变化趋势。

上述推导没有考虑到源、漏空间电荷区的区别,所以等式3只是在 V_{DS} 很小的情况下成立。因为 V_{DS} 一般比源-衬偏置 V_{SB} 大,所以短沟器件的阈值电压更加依赖于 V_D 的大小。更细致的推导应当包含这一效应。

如果MOSFET的沟道宽度减小到与空间电荷区宽度 x_{dmax} 相比拟时,阈值电压也会受到 W 的影响,该效应称为窄沟效应,强烈依赖于制造MOSFET的具体的隔离工艺。在抬高场氧或半凹陷LOCOS隔离技术中(见图9.15),存在两种效应使得阈值电压 V_T 随着 W 的下降而升高。

首先,由于宽度方向边缘电场的作用,部分由栅压感应出来的空间电荷分布在沟道以外的

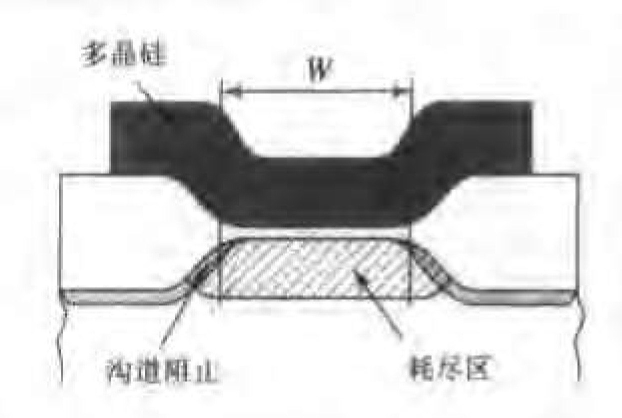


图9.15 半凹陷LOCOS隔离技术的窄沟效应示意图

区域,这样就需要加大栅压来产生足够的感应电荷以形成导电沟道,因而阈值电压升高。第二个原因是隔离区的掺杂浓度大于沟道区。n沟道MOSFET通常需要在器件有源区之间的场区注入杂质硼以防止场区反型。在其后的工艺步骤中,这些场区杂质会扩散到栅极下方的区域,使得沟道边缘的杂质浓度增大。在形成p沟道MOSFET的场氧过程中,磷会自动的在硅表面聚积,在场区形成高浓度的n型杂质。无论上述哪种状况,沟道的边缘杂质浓度都大于沟道的中心部分,这样就提高了阈值电压。在实践中,第二种原因比第一种更重要,特别是那些需要注入高浓度杂质防止场区反型的器件。文献[5]分析了半凹陷MOSFET的这两种效应。更精确的计算需要用数值模拟。

新的工艺常用全凹陷LOCOS隔离技术和沟槽隔离技术,因而出现了第三种与沟道宽度有关的效应,该效应可以降低V_T。因为这个效应与前面的窄沟效应的作用相反,有时被称为反窄沟效应。图9.16是沟槽隔离晶体管的沟道、沟槽场氧化层和栅极的边缘横截面示意图[6]。当栅极加上偏压时,由于沟道两侧的栅电极覆盖了部分场氧化层,来自于重叠区域的电力线聚

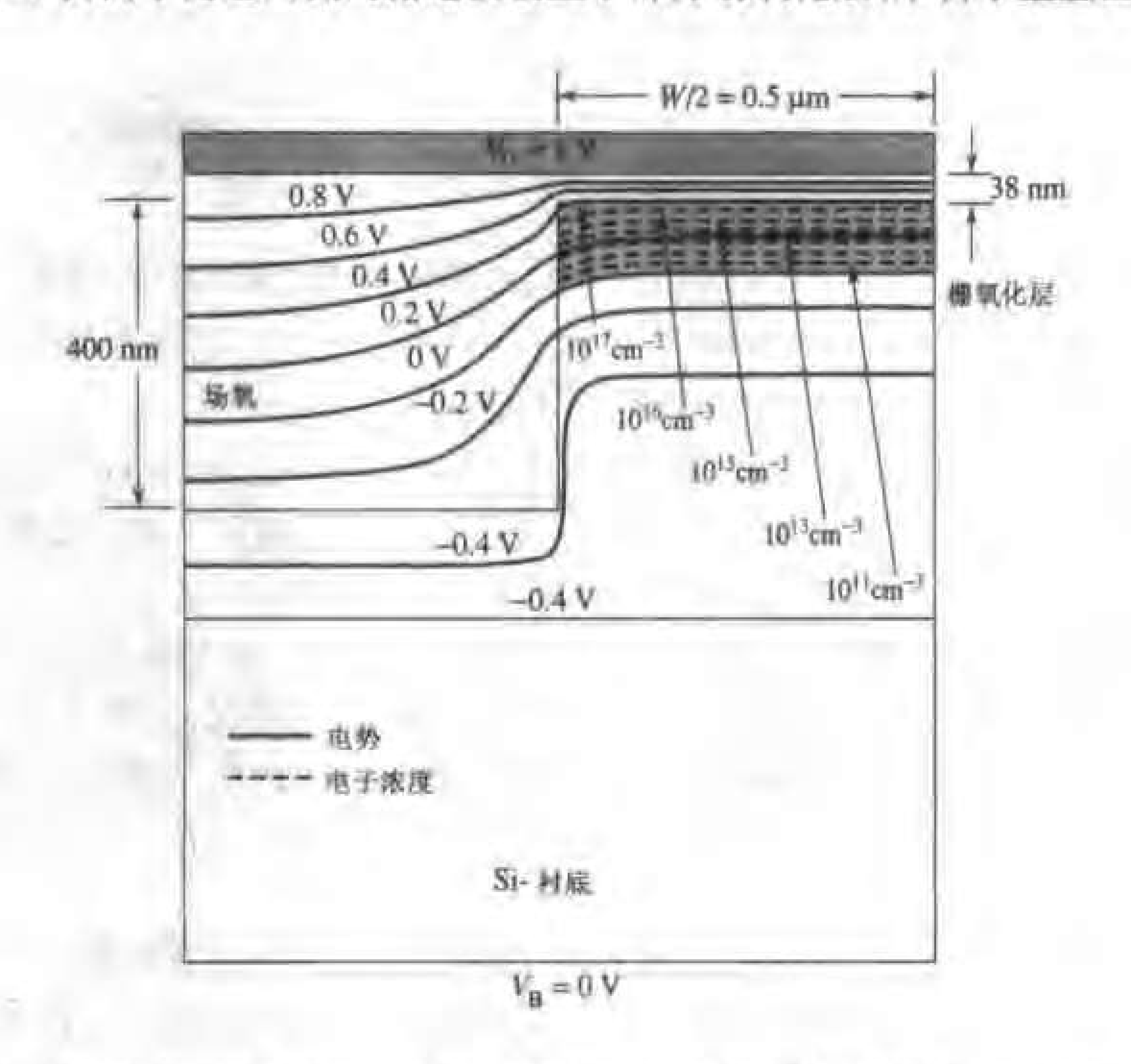


图9.16 全凹陷沟槽隔离MOSFET的二维等位线分部和电子浓度分布(© 1982 IEEE[6])

集在沟道边缘,因而在较低的电压下,就能够在沟道的边缘形成反型层,较小的栅极电压就能够使沿沟道宽度方向的整个沟道区反型。文献[7]给出了理想的沟槽隔离晶体管的解析表达式。不过,反窄沟效应还和其他因素有关[8],包括:(1)硅侧壁的掺杂浓度,影响它的因素包括场区杂质的再分布,生长场氧化层时的杂质分凝,大角度注入工艺导致的侧壁掺杂;(2)沟槽间距;(3)栅边缘拐角区的形状。因此,精确地分析反窄沟效应只能利用数值模拟(假设所有的结构参数已知)

漏致势垒降低 漏致势垒降低(DIBL)指的是漏极电压对源端靠近氧化层表面的 np 结处电子的势垒高度 ϕ_s的影响。如果电子进入沟道的势垒线性下降,从源区进入沟道的电子浓度将指数增长。长沟理论认为只有栅源电压 V_{GS}才能降低源端势垒,但是对于短沟器件,足够高的漏极电压也会降低该势垒,导致阈值电压随 V_D改变,如同随 V_G改变一样,如图 9.17(a)所示。亚阈值电流对 DIBL 效应更为敏感,通过测量亚阈值电流和 V_D的关系很容易检测是否发生了 DIBL 效应。与此相反,长沟器件的亚阈值电流与 V_D 无关。如图 9.17(b)所示,DIBL 效应能大幅度提高亚阈值电流,从而降低了短沟 MOSFET 的性能[9]。

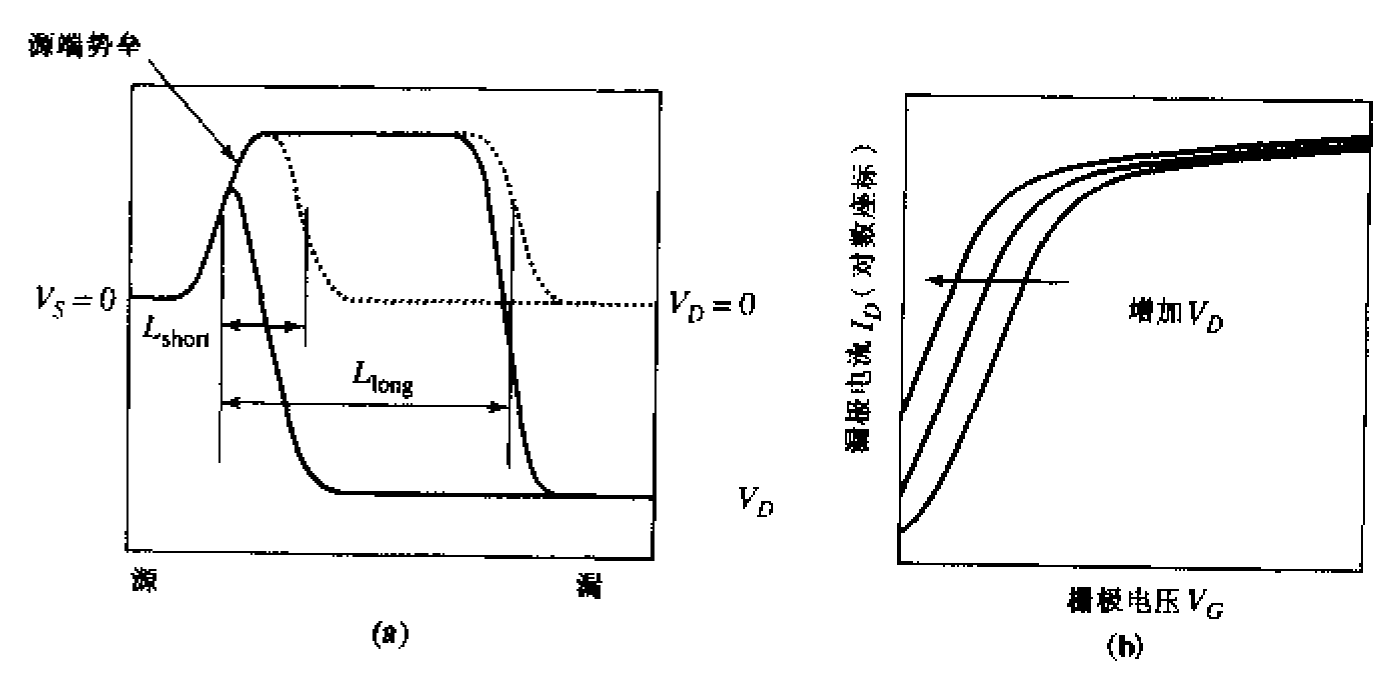


图 9.17 (a) $V_D=0V$(虚线)和高 V_D(实线)时的长沟器件和短沟器件导带图。从图中可看出,高 V_D下短沟器件的源端电子势垒下降;(b)三个不同的 V_D偏压下发生了 DIBL 效应的 MOSFET 的亚阈值导电特性

次表面穿通 类似 DIBL 效应,次表面穿通也是指漏极电压对源端 np 结处电子势垒的影响。与 DIBL 效应不同的是,次表面穿通发生在远离表面的衬底区。次表面穿通的物理机构与 5.6 节讨论的集成电路二极管的穿通原理相同。短沟 NMOSFET 的 p 型表面区的掺杂浓度远大于体掺杂浓度,因此结耗尽区宽度在体内比沟道区大,因此,在足够高的漏极电压之下,就有可能发生漏-衬耗尽区和源-衬耗尽区相互靠近而连接在一起状况,如图 9.18 所示。这一效应能够产生一个穿通区,而且提高漏极偏置会降低穿通区源衬势垒,导致不需要的、通常具有破坏性的漏极电流。图 9.19 绘出了出现次表面穿通的 MOSFET 的 $I_D \sim V_D$曲线和亚阈值区特性曲线。

DIBL 和次表面穿通都是造成集成电路性能下降的原因,最为显著的影响是关态晶体管的泄漏电流较大并且不稳定。这种效应会随着沟道长度的缩短和器件密度的增大而恶化。为了抑制这些效应,设计者通常都尽可能地提高衬底掺杂浓度,减薄栅氧化层厚度和减小源漏结深。

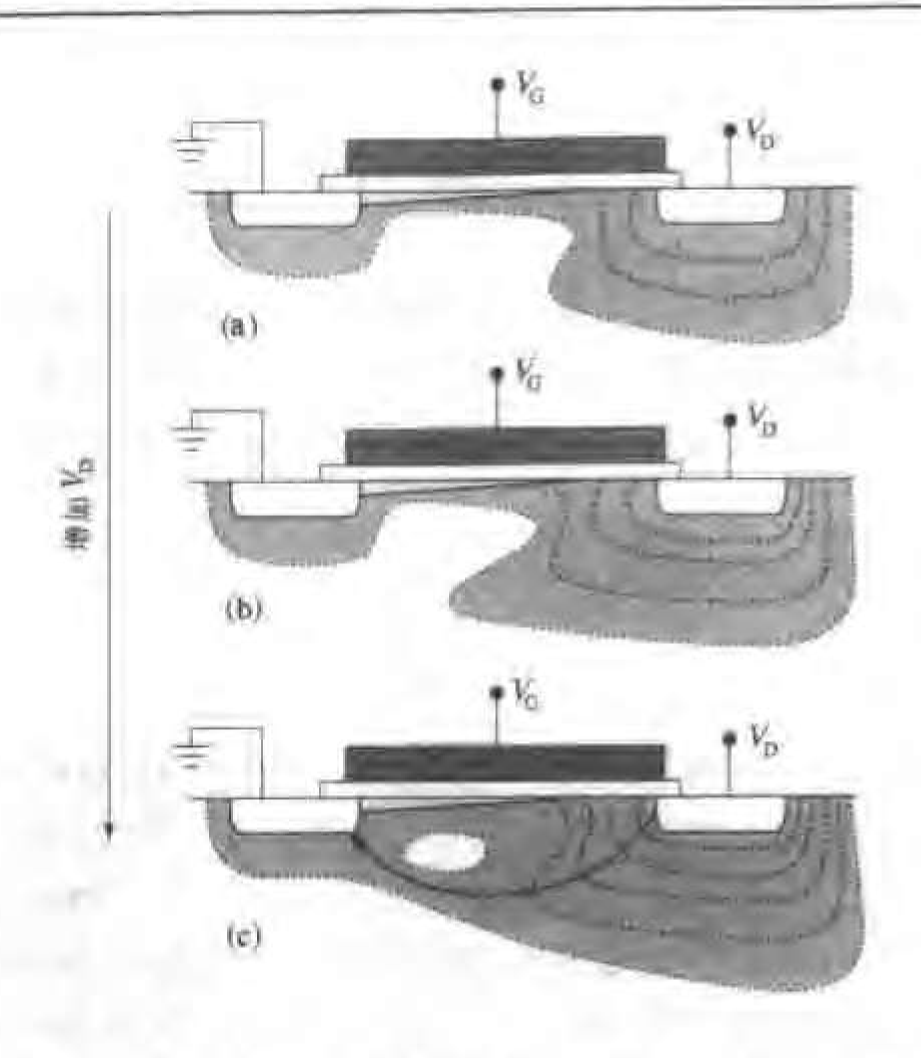


图9.18　短沟MOSFET横截面示意图。MOSFET源极接地，栅极电压$V_G>V_T$并保持不变，漏极电压V_D从图(a)到图(c)不断增大。阴影表示空间电荷区，点线表示等位线。图(c)发生了次表面穿通，实线箭头表示次表面穿通产生的漏极电流通路

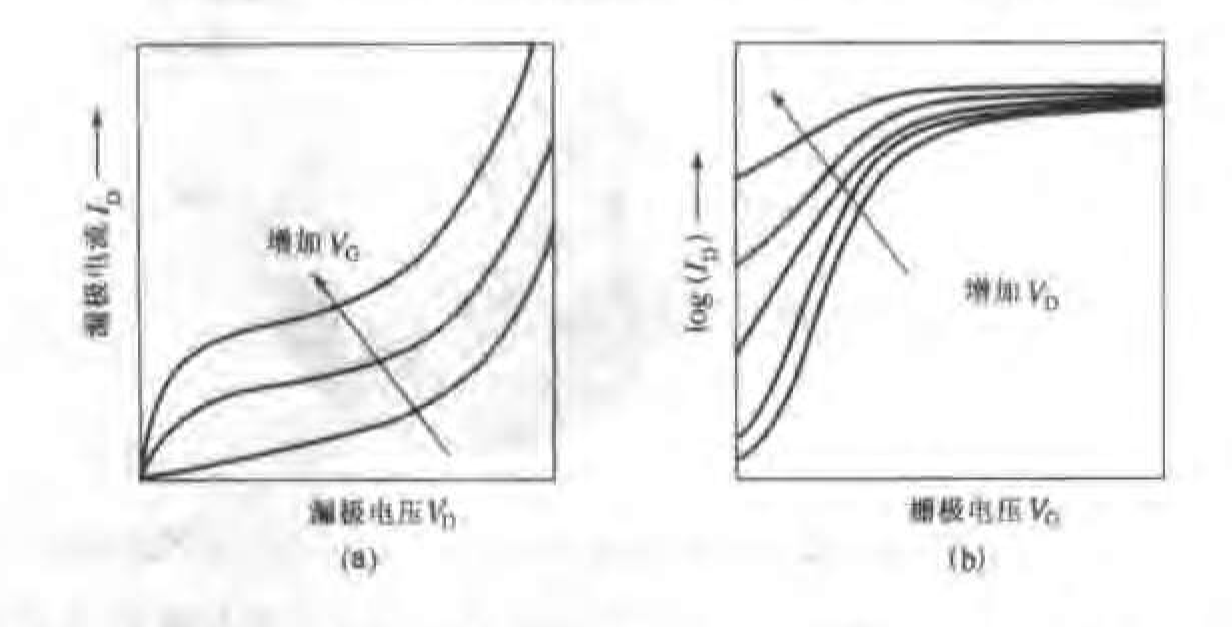


图9.19　MOSFET出现次表面穿通时的(a) I_D-V_D曲线和(b)亚阈值区导电特性

9.2.3　迁移率下降

MOSFET的漏极电流是可动电荷Q_n在电场的作用下，沿着靠近表面区域移动的结果。这些电子受到表面散射、受主杂质散射以及热声子散射的影响[10]。严格求解这些现象对载流子迁移率的影响是不可行的，但是，通过引入表面附近x轴方向的有效电场$\mathscr{E}_{eff}$来作近似处理，可以得到比较简单的迁移率模型[11]。文献[12,13]中的实验证明，当栅氧厚度为4nm时，该迁移率模型对电子和空穴都是成立的。有效电场强度定义为MOS表面所有可动载流子承受的

电场强度的平均值

$$\mathscr{E}_{eff}=\int_0^{\infty}n(x)\mathscr{E}_x(x)\mathrm{d}x\bigg/\int_0^{\infty}n(x)\mathrm{d}x \tag{9.2.1}$$

利用微积分中的公式与8.3节表面总电荷密度Q_s的定义(表面空间电荷区电离受主电荷浓度Q_d与可动电子浓度Q_n之和),推导出$\mathscr{E}_{eff}$的比较简单并且实用的表达式。等式(8.3.14)中,$Q_s=Q_d+Q_n=-\epsilon_s\mathscr{E}_{s0}$,其中$\mathscr{E}_{s0}$是氧化层-硅界面的电场强度。可动电荷浓度$Q_n$是$-qn(x)$沿$x$方向的积分,$Q_n=-q\int_0^{\infty}n(x)\mathrm{d}x$,于是等式(9.2.1)可以写成

$$\mathscr{E}_{eff}=-\frac{1}{\epsilon_s}\left(Q_d+\frac{Q_n}{2}\right) \tag{9.2.2}$$

为了更好地理解等式(9.2.2),我们考虑图9.20所示的一个电子在MOS Si-SiO$_2$界面和表面耗尽区之间的移动。该区域的电力线中止于电离受主或者可动电子上,因为大多数可动电子都在靠近氧化层的表面运动,本质上它们都受到中止于电离受主的电力线影响,也就是等式(9.2.2)中的Q_d项。同时等式(9.2.2)中只有$Q_n/2$出现,原因是中止在表面可动电荷的电力线不会影响其下方运动的载流子,而空间分布的"平均电子"指的是一半在上部,一半是在下部的载流子,所以只有一半的电子对有效场强$\mathscr{E}_{eff}$有贡献。这样,等式(9.2.2)中的$\mathscr{E}_{eff}$就由Q_d和$Q_n/2$决定。

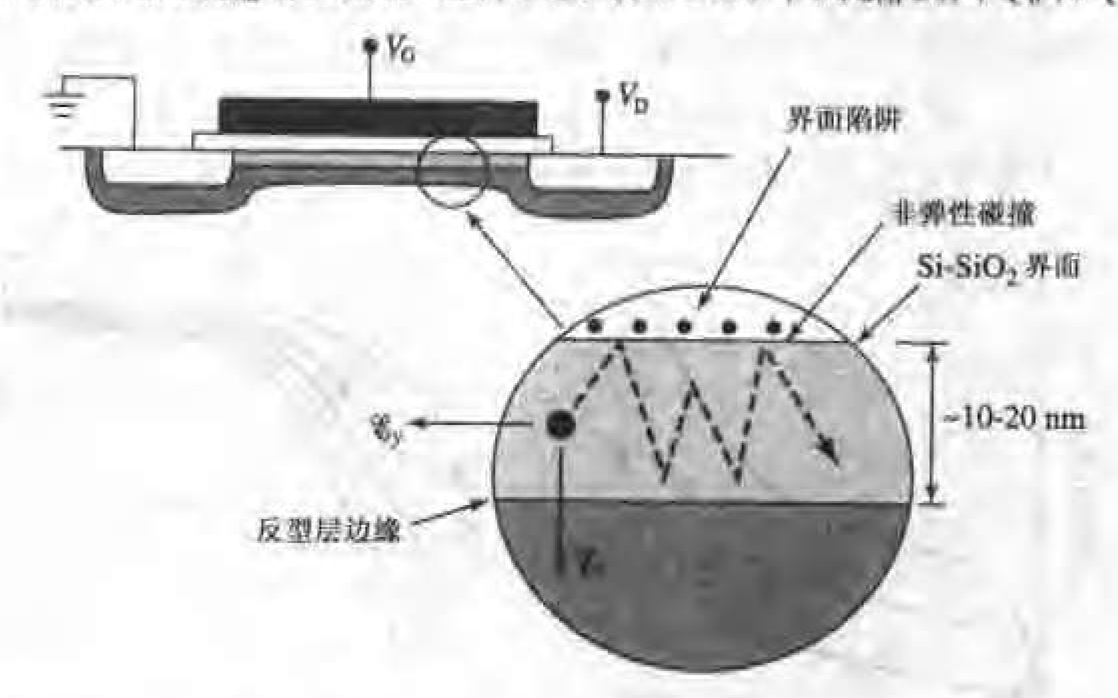


图9.20　在横向电场$\mathscr{E}_y$和纵向电场$\mathscr{E}_x$的作用下,表面反型层电子的运动

等式(9.2.2)还可以进一步推导出更有用的表达式。由8.3节和等式(8.3.19)可以得到$Q_n=-C_{ox}(V_G-V_T)$和$Q_d=-C_{ox}(V_T-V_{FB}-2|\phi_p|)$。为了进一步简化,我们注意到,随着MOSFET尺寸的不断缩小,$(V_{FB}+2|\phi_p|)$基本上保持为一个常数,我们用经验常数V_Z表示,其值约为0.5V。利用V_Z的定义以及Q_d和Q_n的表达式,并考虑到$\epsilon_s/\epsilon_{ox}=3$,由等式(9.2.2)得

$$\mathscr{E}_{eff}=\frac{(V_G-V_T)}{6x_{ox}}+\frac{(V_T+V_Z)}{3x_{ox}} \tag{9.2.3}$$

图9.21是测量得到的n沟道MOSFET电子和p沟道MOSFET空穴的有效迁移率μ_{eff}与x方向表面有效电场$\mathscr{E}_{eff}$(等式(9.2.3))的关系曲线[13]。反映MOSFET的迁移率下降的经验公式为

$$\mu_{eff}=\frac{\mu_0}{1+(\mathscr{E}_{eff}/\mathscr{E}_0)^{\nu}} \tag{9.2.4}$$

式中 μ_0、$\mathscr{E}_0$ 和 ν 都是拟合参数。通过拟合大量的 MOSFET 得出的这些参数值见表 9.3[13]。等式(9.2.4)在较大的范围内都是成立的，在这个范围内载流子主要受声子散射的限制。但是当其他散射机构占支配地位，比如在低温(77K)下库仑散射占支配地位时，这个等式不再成立[14]。

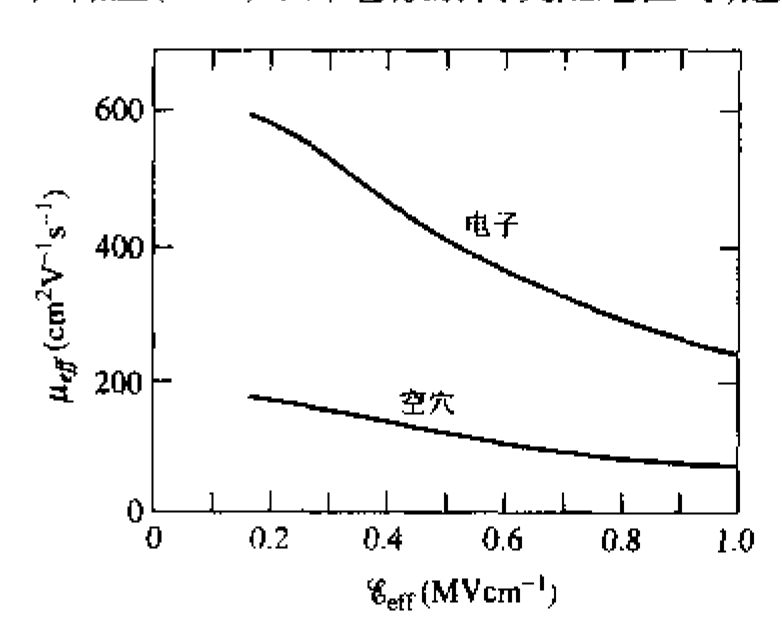


图 9.21　MOSFET 沟道中的电子和空穴的有效迁移率 μ_{eff} 是 x 方向表面有效电场 $\mathscr{E}_{eff}$ 的函数(© 1986 IEEE[13])

表 9.3　有效迁移率参数(© 1986 IEEE[13])

	单　位	电子(表面)	空穴(表面)	空穴(衬底)
μ_0	cm²/V-s	670	160	290
$\mathscr{E}_0$	MV/cm	0.67	0.7	0.35
ν		1.6	1.0	1.0

尽管迁移率下降效应同样影响长沟器件，但是在短沟器件中更明显。当器件按比例缩小时，栅氧厚度按比例缩小的同时电源电压并没有同比例缩小，因此在短沟器件中 x 方向的栅氧化层电场强度更高。

9.2.4　速度饱和

我们在 1.2 节讨论固体中的可动载流子的速度时指出，在弱场条件下，载流子的速度与电场强度呈线性关系，迁移率沿载流子的漂移方向(沿 MOSFET 沟道的 y 轴方向)为常数。强场情形下，载流子漂移速度接近一个极限值(见图 1.18)。MOSFET 的沟道中，因为存在表面散射使得输运受限，弱场下的迁移率也比较小，但是能观测到相似的速度饱和效应。文献[15,16,17]中的实际测量数据指出，表面沟道电子的饱和漂移速度为 $6\sim10\times10^6$ cm/s，空穴的饱和漂移速度为 $4\sim8\times10^6$ cm/s，略低于体内载流子的饱和漂移速度，如图 1.18 所示。测出的表面速度可以用一个较为方便的分段连续的数学表达式拟合，该表达式在速度饱和点(对应于饱和电场 $\mathscr{E}_{sat}$)处导数不连续，电场大于 $\mathscr{E}_{sat}$ 时漂移速度为常数[18]。弱场下的速度可以用下式表示

$$v = \frac{\mu_{\text{eff}}\mathscr{E}}{(1 + \mathscr{E}/\mathscr{E}_{\text{sat}})} \qquad \mathscr{E} < \mathscr{E}_{\text{sat}} \tag{9.2.5}$$

当电场强度大于饱和电场$\mathscr{E}_{sat}$时,载流子速度为常数

$$v = v_{\text{sat}} \qquad \mathscr{E} > \mathscr{E}_{\text{sat}} \tag{9.2.6}$$

在 $v = v_{sat}$处利用等式(9.2.5)求出饱和电场$\mathscr{E}_{sat}$,这样 $v \sim \mathscr{E}$关系是连续的

$$\mathscr{E}_{\text{sat}} = \frac{2v_{\text{sat}}}{\mu_{\text{eff}}} \tag{9.2.7}$$

图9.22为测量数据与等式(9.2.5)和(9.2.6)的拟合曲线。尽管这两个等式存在很多的简化,并且是建立在经验拟合的基础上的,但实践证明它们能较好地预测短沟 MOSFET 的漏极电流。

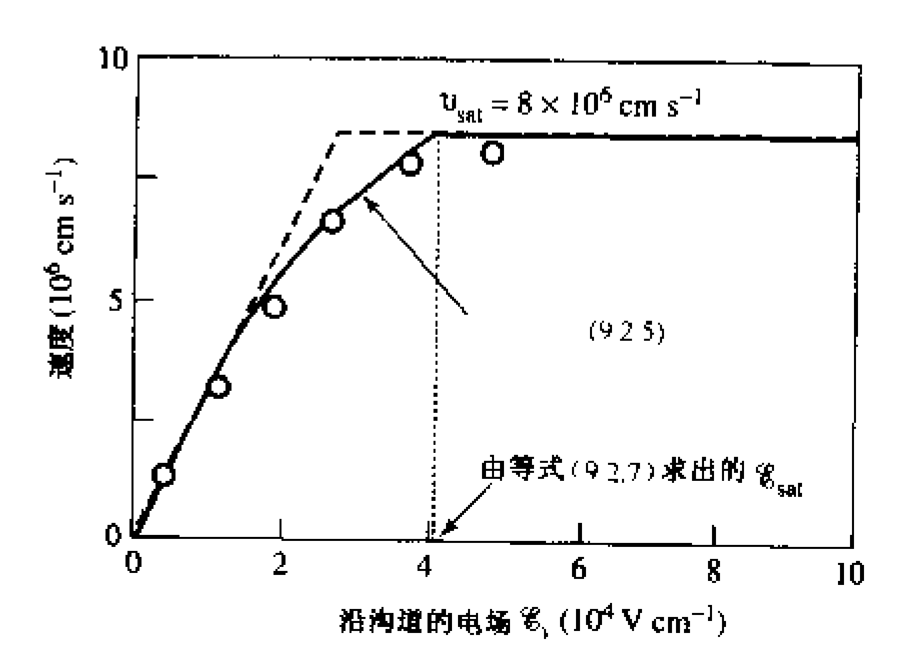


图9.22 在不同横向电场$\mathscr{E}_y$下测量得到的电子漂移速度与等式(9.2.5)的拟合(© 1984 *IEEE*[18])

9.2.5 短沟 MOSFET 的漏极电流

如果考虑等式(9.2.5)中迁移率随横向电场的改变,那么漏极电流表达式(9.1.1)可改写成

$$I_D = C_{ox}W[V_G - V_T - V(y)]\frac{\mu_{\text{eff}}\mathscr{E}(y)}{1 + [\mathscr{E}(y)/\mathscr{E}_{\text{sat}}]} \tag{9.2.8}$$

式中$\mathscr{E}(y) = -\partial V/\partial y \approx -dV/dy$,将$\mathscr{E}(y)$代入等式(9.2.8)得到关于 V 的微分表达式。从源端($V=0$)到漏端($V=V_D$)积分,并假设 V_T是常数,不考虑 Q_d沿沟道方向的变化,得到

$$I_D = \mu_{\text{eff}}C_{ox}\frac{W}{L}\left(V_G - V_T - \frac{V_D}{2}\right)V_D\frac{1}{1 + (V_D/\mathscr{E}_{\text{sat}}L)} \tag{9.2.9}$$

当$\mathscr{E}_{sat} \gg V_D/L$,并且$\mu_{eff} = \mu_n$时,等式(9.2.9)与等式(9.1.5)相同。等式(9.2.9)的推导将μ_{eff}看做常数,尽管等式(9.2.4)已经指出μ_{eff}与表面有效电场$\mathscr{E}_{eff}$有关,沿 y 方向的变化明显。另外,等式(9.2.9)还忽略了 y 方向 V_T的变化(由于忽略了体电荷的变化)。即使存在这些近似,等式(9.2.9)还是足够精确的,并且非常简单实用。

漏极电流饱和 由于载流子速度具有上限,当载流子以饱和漂移速度 v_{sat} 到达漏极时,漏极电流也会饱和。如果我们将此时的漏极电压定义为饱和漏极电压,则漏端单位面积的电子电荷量 Q_{nD} 可以表示为 $Q_{nD} = -WC_{ox}(V_G - V_T - V_{Dsat})$。这些电荷移动速度为 v_{sat},则漏极电流表达式为

$$I_{Dsat} = WC_{ox}(V_G - V_T - V_{Dsat})v_{sat} \tag{9.2.10}$$

将等式(9.2.10)代入等式(9.2.11),求出饱和漏极电压 V_{Dsat}

$$V_{Dsat} = \frac{\mathscr{E}_{sat}L(V_G - V_T)}{\mathscr{E}_{sat}L + (V_G - V_T)} \tag{9.2.11}$$

正如我们所预期的,如果饱和电场 $\mathscr{E}_{sat}$ 非常大,可以不考虑速度饱和效应,那么等式(9.2.11)近似为 $(V_G - V_T)$。

下面举例说明短沟模型的应用。

例题 短沟模型

用短沟模型计算 n 沟道 MOSFET 的 V_{Dsat} 和 I_{Dsat}。已知 $x_{ox} = 20\text{nm}$,$W = 50\mu\text{m}$,$L = 0.5\mu\text{m}$,$V_T = 0.7\text{V}$,$V_S = 0$,$V_G = 3\text{V}$,$V_D = 1.5\text{V}$。将计算结果和长沟模型进行比较。

解:取 $V_Z = 0.5\text{V}$,由等式(9.2.3)得

$$\mathscr{E}_{eff} = \frac{(V_G - V_T)}{6x_{ox}} + \frac{(V_T + V_Z)}{3x_{ox}} = 3.92 \times 10^5\ \text{Vcm}^{-1}$$

利用等式(9.2.4)和表9.3

$$\mu_{eff} = \frac{\mu_0}{1 + (\mathscr{E}_{eff}/\mathscr{E}_0)^{\nu}} = 470\ \text{cm}^2\text{V}^{-1}\text{s}^{-1}$$

取 $v_{sat} = 8 \times 10^6\text{cm/s}$,计算出饱和电场 $\mathscr{E}_{sat} = 2v_{sat}/\mu_{eff} = 3.4 \times 10^4\text{V/cm}$

由等式(9.2.11)求出 $V_{Dsat} = \dfrac{\mathscr{E}_{sat}L(V_G - V_T)}{\mathscr{E}_{sat}L + (V_G - V_T)} = 0.98\text{V}$

因为漏极电压为1.5V($>V_{Dsat}$),MOSFET 工作在漏极电流饱和区,利用等式(9.2.10)得

$$I_{Dsat} = WC_{ox}(V_G - V_T - V_{Dsat})v_{sat} = 9\ \text{mA}$$

为了与长沟模型做比较,我们利用长沟理论计算:$V_{Dsat} = V_G - V_T = 2.3\text{V}$,MOSFET 工作在线性区,由等式(9.1.5)得

$$I_D = \mu_n C_{ox}\frac{W}{L}\left[\left(V_G - V_T - \frac{1}{2}V_D\right)V_D\right] \quad 和 \quad I_D = 24\ \text{mA}$$

式中低场下的迁移率取典型值 $600\text{cm}^2\text{V}^{-1}\text{s}^{-1}$。长沟理论和短沟理论得到结果相差很大,说明了短沟效应对 MOSFET 的性能影响很大。

跨导 当 MOSFET 用于放大器时,饱和区跨导 $g_{msat} = \partial I_{Dsat}/\partial V_G$ 是非常重要的参数。在长沟模型的等式(9.1.6)中对 V_G 求导,可以得出饱和区跨导 g_{msat}

$$g_{msat} = \frac{2I_{Dsat}}{(V_G - V_T)} \tag{9.2.12}$$

考虑速度饱和，利用等式(9.2.10)计算得到 g_{msat}

$$g_{msat} = Wv_{sat}C_{ox}\left[1 - \frac{\partial V_{Dsat}}{\partial V_G}\right] \tag{9.2.13}$$

从等式(9.2.11)和(9.2.7)可以知道 V_{Dsat}和 V_G的关系，因此短沟 MOSFET 的 g_{msat}表达式为

$$g_{msat} = Wv_{sat}C_{ox}\frac{(V_G - V_T)(V_G - V_T + 2\mathscr{E}_{sat}L)}{(V_G - V_T + \mathscr{E}_{sat}L)^2} \tag{9.2.14}$$

等式(9.2.9)，(9.2.10)，(9.2.11)和(9.2.14)组成了最基本的描述短沟 MOSFET 特性的方程组。

9.2.6 MOSFET 的按比例缩小和短沟模型

虽然前面讨论的短沟 MOSFET 模型公式从形式上看相对比较简单，但对 MOSFET 集成电路的发展是非常有用的，比如对 MOSFET 的按比例缩小具有指导作用。下面我们讨论 MOSFET 按比例缩小会遇到的各种效应，并进一步发展短沟模型。对这些问题的考虑有助于我们从物理上理解短沟效应。

非饱和区的漏极电流 比较等式(9.1.5)和(9.2.9)可知，短沟效应对非饱和区漏极电流的影响比较小。在低漏极偏置区域，导致漏极电流下降的主要原因是迁移率下降。随着栅极电压 V_G的升高，迁移率下降的影响更为明显(等式(9.2.3)和(9.2.4))。

饱和漏极电压 从等式(9.2.11)可以看出，饱和漏极电压 V_{Dsat}不仅与有效栅极偏压($V_G - V_T$)有关，还与饱和电场$\mathscr{E}_{sat}$和沟道长度 L 的乘积有关。当$\mathscr{E}_{sat}L \gg (V_G - V_T)$时，饱和漏极电压趋于长沟 MOSFET 的饱和漏极电压($V_G - V_T$)。随着 MOSFET 尺寸的不断缩小，饱和漏极电压趋于 $V_{Dsat} = \mathscr{E}_{sat}L$。MOSFET 的按比例缩小要求栅氧化层厚度 x_{ox}也按比例缩小，对应的氧化层电容 C_{ox}就要增大，表面场强也随之增大，导致表面迁移率μ_{eff}下降(等式(9.2.7))，$\mathscr{E}_{sat}$和 V_{Dsat}提高。

饱和漏极电流 在长沟理论中，饱和漏极电流与有效栅电压($V_G - V_T$)之间是平方关系[$I_{Dsat} \propto (V_G - V_T)^2$](等式(9.1.6))，但是在短沟理论中是线性关系[$I_{Dsat} \propto (V_G - V_T - V_{Dsat})$](等式(9.2.10))。图 9.23 绘出了两种理论计算得出的 MOSFET 特性的比较，二者区别显著。可能我们会觉得等式(9.2.10)有些奇怪，因为虽然是从短沟效应推导出来的表达式，但是并没有 L 出现在表达式中。L 对 I_{Dsat}的影响是通过对饱和漏极电压 V_{Dsat}的影响来体现的，L 下降，V_{Dsat}也会下降(等式(9.2.11))。从物理上看，随着 L 的减小，可动电荷 Q_n中以饱和速度 v_{sat}漂移的部分增大，导致 I_{Dsat}增大。如果 L 足够小，从等式(9.2.11)[2]可看出 V_{Dsat}趋于零，但是从等式(9.2.10)可得出最大饱和漏极电流为

$$I_{Dmax} = v_{sat}WC_{ox}(V_G - V_T) \tag{9.2.15}$$

与之不同的是，长沟理论(等式(9.1.6))预示当 L 趋于零时，饱和漏极电流趋于无穷大($I_{Dsat} \propto L^{-1}$)。图 9.24 比较了这两种不同理论的计算结果。

---

[2] 译者注：原文误为等式(9.1.11)。

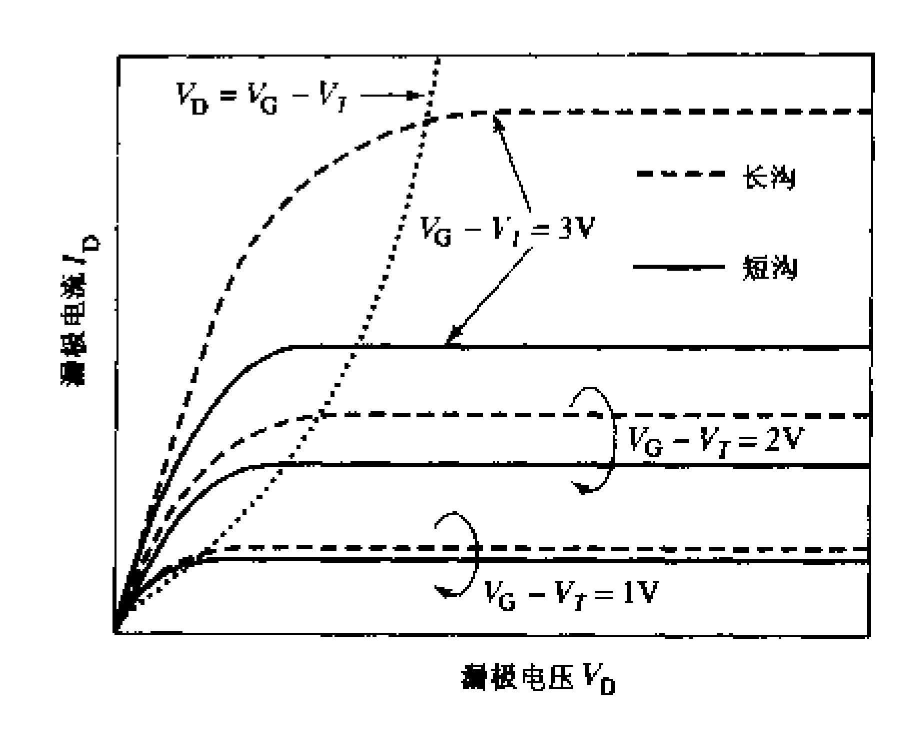


图 9.23　长沟理论（虚线）和短沟理论（实线）所预示的漏极电流特性

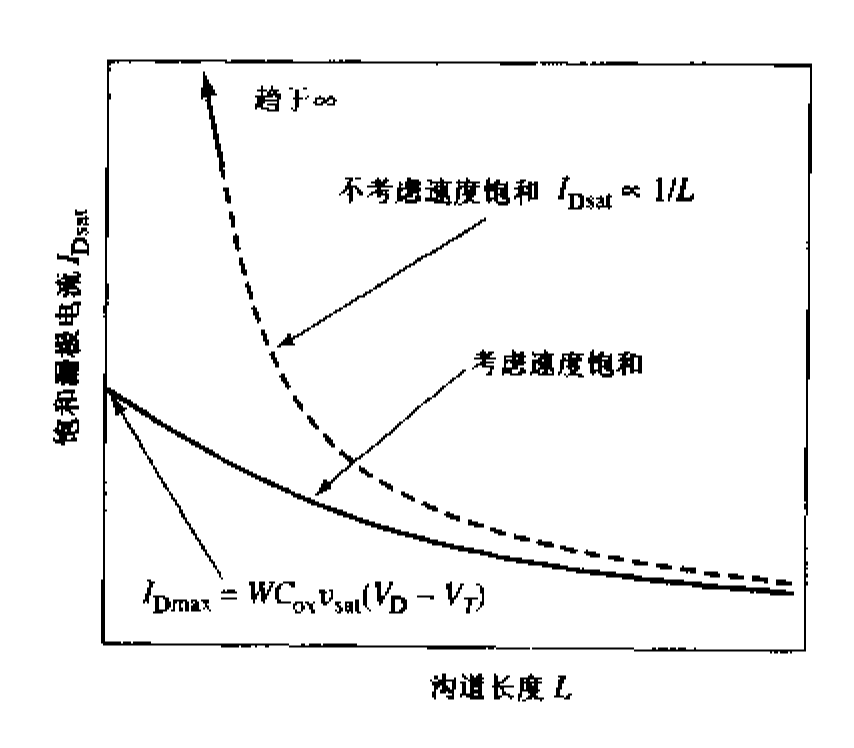


图 9.24　MOSFET 饱和漏极电流与沟道长度的关系。
虚线不考虑速度饱和，实线考虑速度饱和

等式(9.2.15)计算出的 I_{Dmax} 在 MOSFET 按比例缩小的时候是一个很有用的标志。我们可以定义按比例缩小 MOSFET 的理想因子 K_I，其值为按比例缩小后的 I_{Dsat} 与极限参数 I_{Dmax} 的比值。利用等式(9.2.10)和(9.2.15)得到 K_I 的表达式

$$K_I = \frac{I_{Dsat}}{I_{Dmax}} = \frac{(V_G - V_T - V_{Dsat})}{(V_G - V_T)} \tag{9.2.16}$$

很容易看出 K_I 总是小于 1；不明确的是 K_I 值与 MOSFET 制造工艺的关系。例如，$x_{ox} = 40\text{nm}$，$L = 1\mu\text{m}$ 的 MOSFET，$(V_G - V_T) = 4.3\text{V}$ 时，$V_{Dsat} = 1.3\text{V}$，理想因子 $K_I = (4.3 - 1.3)/4.3 = 0.72$。这表明进一步降低器件的沟道长度，最好的情况是能将器件的驱动电流提高约 40%。图 9.25 绘出了两种栅氧厚度（7nm 和 20nm）条件下，理想因子 K_I 随着 L 的变化。从图中可以看出，薄栅氧器件的 K_I 随 L 的变化更为明显，因此减小 MOSFET 的尺寸对薄栅氧器件的漏极电流影响更大，并且随着沟道长度的缩短，这种影响将更加明显。薄栅氧器件由于栅电容较大，栅电容趋向于在所有效应中占主导地位，因此 MOSFET 的设计非常重视减小 x_{ox}，从而增大电容 C_{ox}。

沟道非常短的 MOSFET 的测量数据(图 9.26[19])证明了上述结论是正确的。但是过度减薄栅氧会带来可靠性和成品率的问题,需要折衷考虑

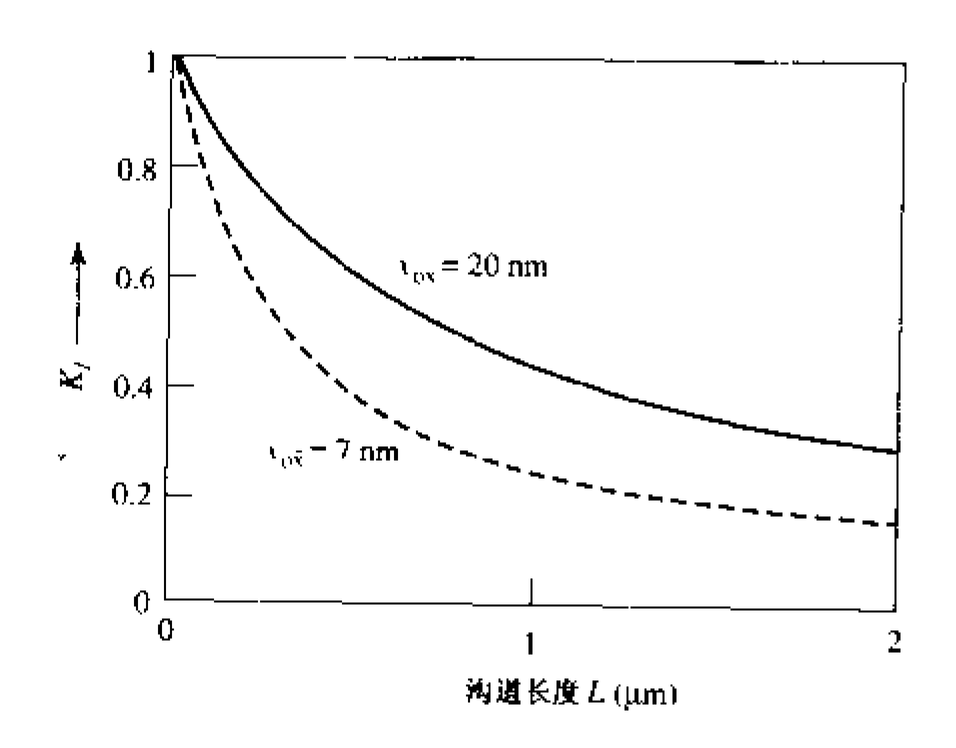


图 9.25 不同的栅氧厚度下计算得到的漏极电流理想因子 K_l 与沟道长度 L 的关系

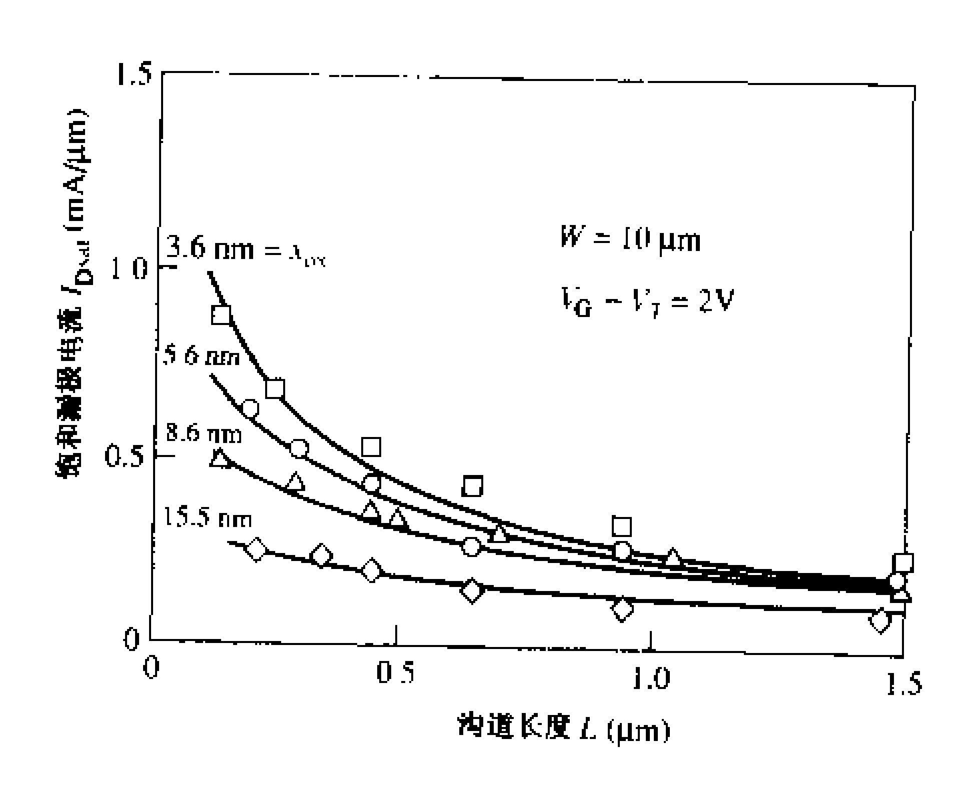


图 9.26 不同栅氧厚度的深亚微米(短沟)n 沟道 MOSFET 的饱和漏极电流 I_{Dsat} 的测量值和计算曲线(© 1987 IEEE[19])

短沟 MOSFET 的工作速度 短沟 MOSFET 的工作速度可以用类似于等式(9.1.34)的推导方法求解,但是得到的理论表达式过于复杂,没有任何指导意义。短沟 MOSFET 的 L' 到 L 之间的速度饱和区占整个沟道长度相当大的部分(见图 9.4(c)),速度饱和区的具体求解将在第 10 章介绍,这里只定性分析 MOSFET 的开关速度。

为了表征电子器件的最大开关速度,可以通过测量以频率为参数的器件的放大性能,记录增益降为 1 时的截止频率 f_T。MOSFET 的 f_T 表达式可以从 9.1 节中的小信号模型得出

$$f_T = \frac{g_{msat}}{2\pi C_{in}} \tag{9.2.17}$$

其中 g_{msat} 是 MOSFET 的饱和区跨导，C_{in} 是输入电容。

对于长沟 MOSFET，g_{msat} 可以从等式(9.1.37)得到，C_{in} 主要来源于栅源电容 C_{GS}[饱和区的 $C_{GS}=(2/3)C_{ox}$]，因此，长沟 MOSFET 的 f_T 为

$$f_T=\frac{3\mu_n(V_G-V_T)}{4\pi L^2} \tag{9.2.18}$$

等式(9.1.34)是长沟 MOSFET 的渡越时间 T_{tr} 的表达式，与等式(9.2.18)相比，可以发现 $f_T=1/\pi T_{tr}$。

对于短沟器件，跨导的表达式为等式(9.2.14)，代入等式(9.2.17)，有

$$f_T=\frac{3v_{sat}(V_G-V_T)(V_G-V_T+2\mathscr{E}_{sat}L)}{4\pi L(V_G-V_T+\mathscr{E}_{sat}L)^2} \tag{9.2.19}$$

考虑到短沟限制，$\mathscr{E}_{sat}L \ll (V_G-V_T)$，等式(9.2.19)可以简化成

$$f_T=\frac{3v_{sat}}{4\pi L} \tag{9.2.20}$$

比较等式(9.2.18)和(9.2.20)可以看出，MOSFET 尺寸的减小对于器件的特性有巨大的影响，f_T 对沟道长度 L 的依赖从 L^{-2} 变为到 L^{-1}，如图 9.27 所示。本节推导出的 f_T 的值只与 MOSFET 本身的结构有关，代表了无负载时器件具有放大能力的最高频率。实际应用中，MOSFET 实际的 f_T（或开关速度）主要是由器件几个节点处的电容决定的，比本征器件的 f_T 要小。

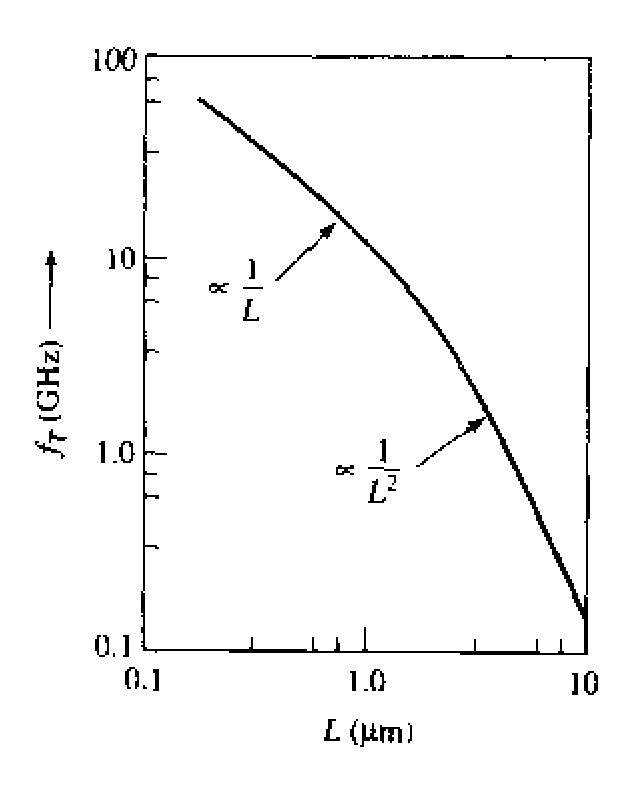


图 9.27　n 沟道 MOSFET 截止频率 f_T 与沟道长度 L 的关系曲线（利用等式(9.2.14)和(9.2.17)计算）

9.3　器件：互补 MOSFET-CMOS

在 MOS IC 发展的早期，人们就发现如果数字电路中将 p 沟道 MOSFET 和 n 沟道 MOSFET 串联，则能够大大减小“静态”(dc)功耗[20]，这种电路被称为互补 MOS 晶体管电路，或者简称 CMOS 电路。最基本的 CMOS 电路模块是反相器电路，它的输出是输入的二进制信号的互补

信号。图 9.28 是 CMOS 反相器的电路图、版图和电压传输特性。反相器中的两个器件相互串联(p 沟道 MOSFET 的漏极和 n 沟道 MOSFET 的漏极相连),它们的栅极连在一起。

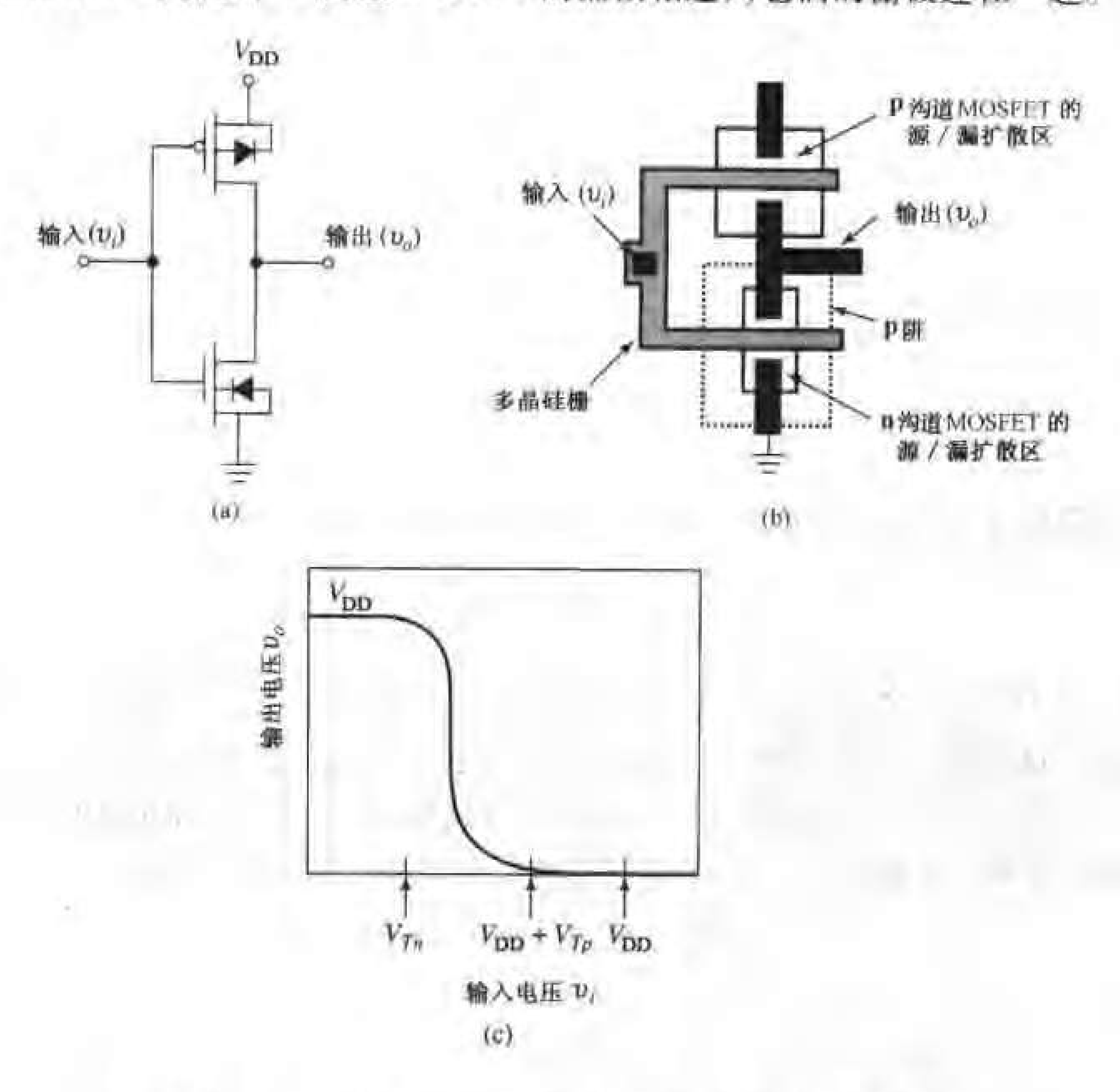


图 9.28　(a) CMOS 反相器电路;(b) p 阱 CMOS 反相器版图;(c) CMOS 反相器的电压传输特性

为了了解反相器的工作原理,先假设输入电压小于 n 沟道 MOSFET 的阈值电压,但是与 p 沟道 MOSFET 的衬底电压相比足够负,因此 n 沟道 MOSFET 处于关态(不导电),p 沟道 MOSFET 处于开态(导电)。在这种情况下,p 沟道 MOSFET 提供从电源到输出端的通路,n 沟道 MOSFET 隔离了输出端到地的通路。通常,输出端接到下一级反相器的输入,没有直流通路,因此输出电压处在"高"状态($v_o \approx V_{DD}$)。如果这时升高输入电压,p 沟道 MOSFET 转为截止。当输入电压大于 n 沟道 MOSFET 的阈值电压时,n 沟道 MOSFET 形成沟道,将输出电压从 V_{DD} 拉到地电平。因此在直流条件下,两个 MOSFET 的其中一个总是处在关态,所以除了器件的泄漏电流以外,没有直接的从正电源到地的电流通路。因为这个原因,CMOS 的电压传输特性非常陡峭,几乎所有的功耗都发生在高低电平转换时,此时需要对输入端和输出端节点充放电。这种陡峭的电压传输特性正是数字电路所需要的,并且该电路还能抵抗电源和地端逻辑信号的噪声干扰。正是由于这些优点,导致 CMOS 在现代 IC 中占主导地位。

9.3.1　CMOS 的设计考虑

因为 CMOS 工艺最基本的思想是在同一个集成电路中同时制造 n 沟道 MOSFET 和 p 沟道

MOSFET,所以其中有一种晶体管必须放置在与衬底导电类型相反的区域中,这样的区域通常是用离子注入和扩散形成的,通常称做"阱"。如果衬底是n型的,那么p沟道MOSFET就直接做在衬底上,同时需要形成p型区(p阱)以制备n沟道MOSFET。类似的,CMOS也可以做在p型衬底n阱上。

n阱工艺和p阱工艺都有各自的优缺点。阱是通过补偿衬底杂质形成的,因此总杂质浓度较高,载流子迁移率较低。在同样的掺杂浓度下,阱中MOSFET的沟道载流子的迁移率通常比衬底上的MOSFET低。因为n沟道MOSFET和p沟道MOSFET要求近似相等的驱动电流,我们更愿意选择p阱工艺(n沟道MOSFET做在p阱中),这是因为电子的迁移率比空穴的迁移率高。但是现代的集成电路并不全是由CMOS反相器构成的,通常某些电路部分只含有n沟道MOSFET,所以在整个IC中n沟道MOSFET的数量大于p沟道MOSFET。在这种情况下,我们更愿意采用图9.29所示的n阱工艺。

在设计CMOS工艺时,为了避免从漏区底部到下方衬底的穿通,需要重点考虑阱深。下面举例说明这一问题。

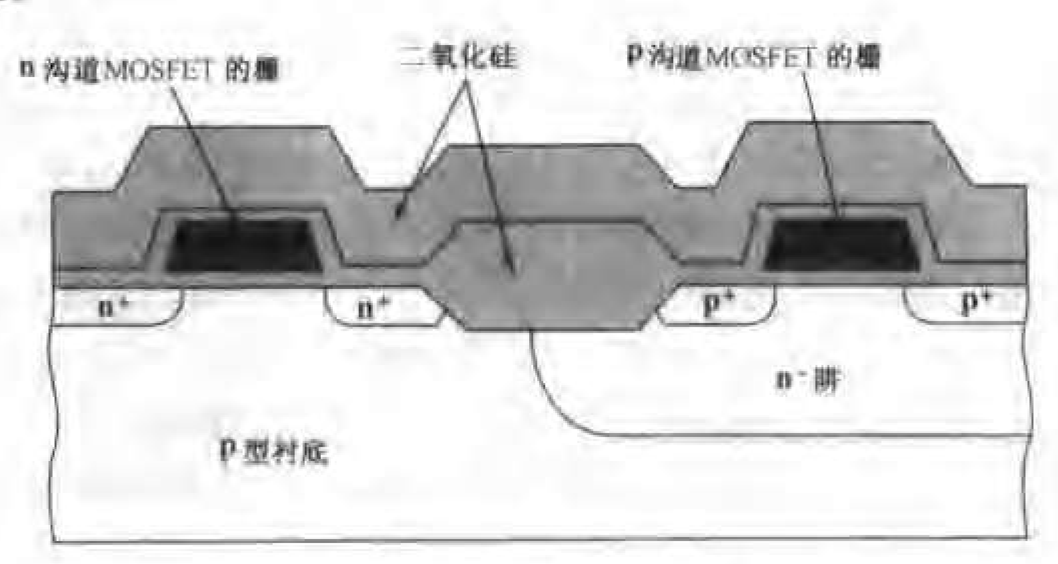


图9.29　n阱CMOS反相器横截面示意图

例题　CMOS阱深设计

已知n阱CMOS工艺,要求电路的工作电压$V_{DD}=1.5\text{V}$,p型硅片掺杂浓度$N_a=5\times10^{14}\text{cm}^{-3}$,n阱的平均掺杂浓度$N_d=3\times10^{15}\text{cm}^{-3}$。p沟道MOSFET的源漏结深$x_j=0.8\mu\text{m}$,平均掺杂浓度$N_a=8\times10^{18}\text{cm}^{-3}$。为了防止与衬底之间的垂直穿通,请问最小阱深为多少?

解:垂直穿通发生在从p沟道MOSFET的源端($V_S=V_{DD}=1.5\text{V}$)到接地的衬底之间的两个背靠背的pn结,如图9.29所示。

源到n阱的pn结是单边突变结,内建电势为$\phi_i\approx0.78\text{V}$,由表4.1和等式(4.3.1)知,n阱一侧的耗尽区宽度为$0.58\mu\text{m}$。阱到衬底的pn结内建电势为$\phi_i\approx0.58\text{V}$,由等式(4.3.1)知,偏压1.5V下的总耗尽区宽度为$2.51\mu\text{m}$。利用等式(4.2.6)得出,其中1/7的耗尽区宽度($0.36\mu\text{m}$)扩展在阱区。因此阱区的最小深度应包括源漏结深$0.8\mu\text{m}$,加上阱区中的耗尽区宽度之和(0.58+0.36),共$1.74\mu\text{m}$。实际预算会比这个值稍大,以保证一定的安全范围。

另外需要考虑的是当p沟道MOSFET处在关态时,它的漏极接地,此时阱到漏区的耗尽区宽度比阱到源区的耗尽区宽度大,因为漏极和阱之间有电压降V_{DD}。然而,即使此时漏

区—阱和阱—衬底两个 pn 结的耗尽区相互连接,也不会产生很大的穿通电流,因为此时漏极和衬底都接地。虽然如此,这样的设计也是不可取的,因为它们很容易造成阱区的其他地方发生穿通。利用和前面类似的分析,为了保证在任何偏压下,阱区内都存在电中性区,可求出阱区的最小深度为 2.16μm。考虑到一定的安全范围,合适的阱深是 2.5μm。

为了制造高性能的 CMOS 电路,现代 CMOS 工艺设计采用了更加精细和昂贵的制造工艺,以保证 n 沟道 MOSFET 和 p 沟道 MOSFET 都能有最好的性能。双阱,甚至三阱 CMOS 工艺已经被采用。图 9.30 为双阱 CMOS 工艺步骤。

9.3.2 MOSFET 参数和参数提取

集成电路设计的基本工具是电路模拟软件,而精确的模拟结果是建立在精确的模型上的为了获得精确的模型,人们做了大量的工作。在众多的 MOSFET 模型中,我们重点讨论由加州大学 Berkeley 分校提出的 BSIM 模型[21]。BSIM3v3(Berkeley Short-channel IGFET Model 3 Version 3)已经被 Compact Model Council(由半导体工业界各领先公司组成的机构)[22]定为工业界的标准,BSIM3v3 模型对于沟道长度短至 0.1μm 的 MOSFET 都有效,因此它包含了许多复杂的物理效应。9.2 节已经描述了 BSIM3v3 的基本物理框架,这部分内容只讨论如何获得 BSIM3v3 的模型参数。参数提取能将电路设计和工艺流程联系起来,因此非常重要。参数提取的第一步是从 IC 制造工厂,例如代工厂获得 CMOS 的原型样品,对其进行测试,从测试数据中提取出 BSIM3v3 的模型参数,不需要参考 CMOS 原型样品的制造流程。还有其他的计算器件参数的方法,比如利用已知的或模拟出的器件结构的更为复杂的模型,不过我们在此就不讨论了。

全局优化 参数提取的任务是从一组器件测试特性中得到与器件模型相对应的一套模型参数,使在所有测试数据范围内,模拟结果与测试数据之间误差最小。我们用参数 ϵ 表示误差,定义均方根误差为

$$\epsilon^2 = \sum_i [D_i|(V^i_{G,D,S,B}) - M_i(p_1, p_2, \dots, p_n)|(V^i_{G,D,S,B})]^2 \tag{9.3.1}$$

式中 D_i 是在偏置($V^i_{G,D,S,B}$)下的测试数据;M_i 是在相同的偏置下的模拟数据,使用的参数为 $\{p_1, p_2, p_3, \cdots, p_n\}$。我们的目标就是寻找一套参数 $\{p_1, p_2, p_3, \cdots, p_n\}$,使等式(9.3.1)中的误差最小。

全局优化将参数提取看做是纯粹的最小化问题,用计算机算法找出能最好地拟合所有的测试数据的一套参数值。因为这种方法是纯经验的,所以用这种方法提取的任何参数都有可能与物理模型计算出来的值有较大的偏离。在某些情况下,提取出的某些参数会失去物理意义。尽管这样,全局优化在提取器件参数的工作区还是很有用的。但是,将这套参数用在其他工作区,可能导致较大的误差。要成功地利用误差最小化方法,必须在较大的偏置范围内进行测试,要求考虑器件在集成电路中可能遇到的所有情况。

局部优化 局部优化指的是对器件按不同的工作区进行参数提取,在指定的工作区提取与这个区域对应的物理机制占主导地位的参数。与全局优化相比,局部优化得到的参数有时不能在所有的工作区都能与测试数据很好地吻合,但它们与器件的物理过程联系紧密。由于

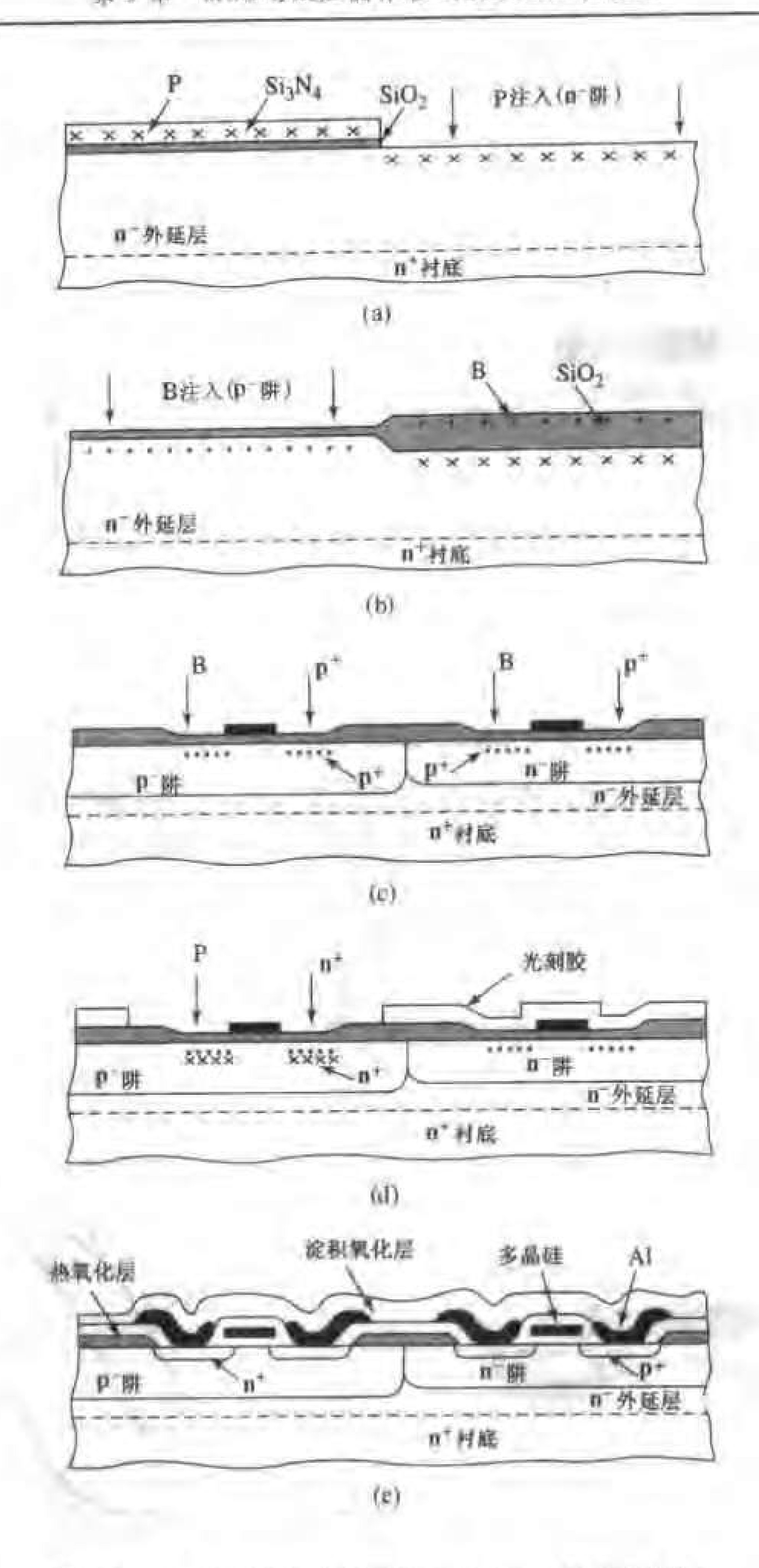


图 9.30　CMOS VLSI 中的双阱工艺(© 1980 IEEE[29])

这个原因,本节将以几个具体的模型参数(包括栅氧化层厚度 x_{ox},低场阈值电压 V_T以及有效迁移率 μ_{eff})的提取为例详细讨论局部优化。

栅氧化层厚度　提取栅氧化层厚度 x_{ox}的电路如图 9.31(a)所示,电路中 MOSFET 源极和漏极都接地,栅极正偏,衬底浮空。当外加的栅电压缓慢上升(准直流)时,测量得到栅极和源/漏之间电容。典型的测量数据如图 9.31(b)所示(与 8.6 节讨论的栅控二极管类似),$V_G = V_T$时,电容发生突变。电容的变化量等于 $C_{ox}WL$,其中 W 和 L 可以从版图获得,这样就能求出 C_{ox},参数 $x_{ox} = \epsilon_{ox}/C_{ox}$也就提取出来了。为了减小其他效应带来的误差,通常在大尺寸器

件上测量,一般 W 和 L 都大于 10μm,以保证 C_{ox} 和 x_{ox} 足够精确。

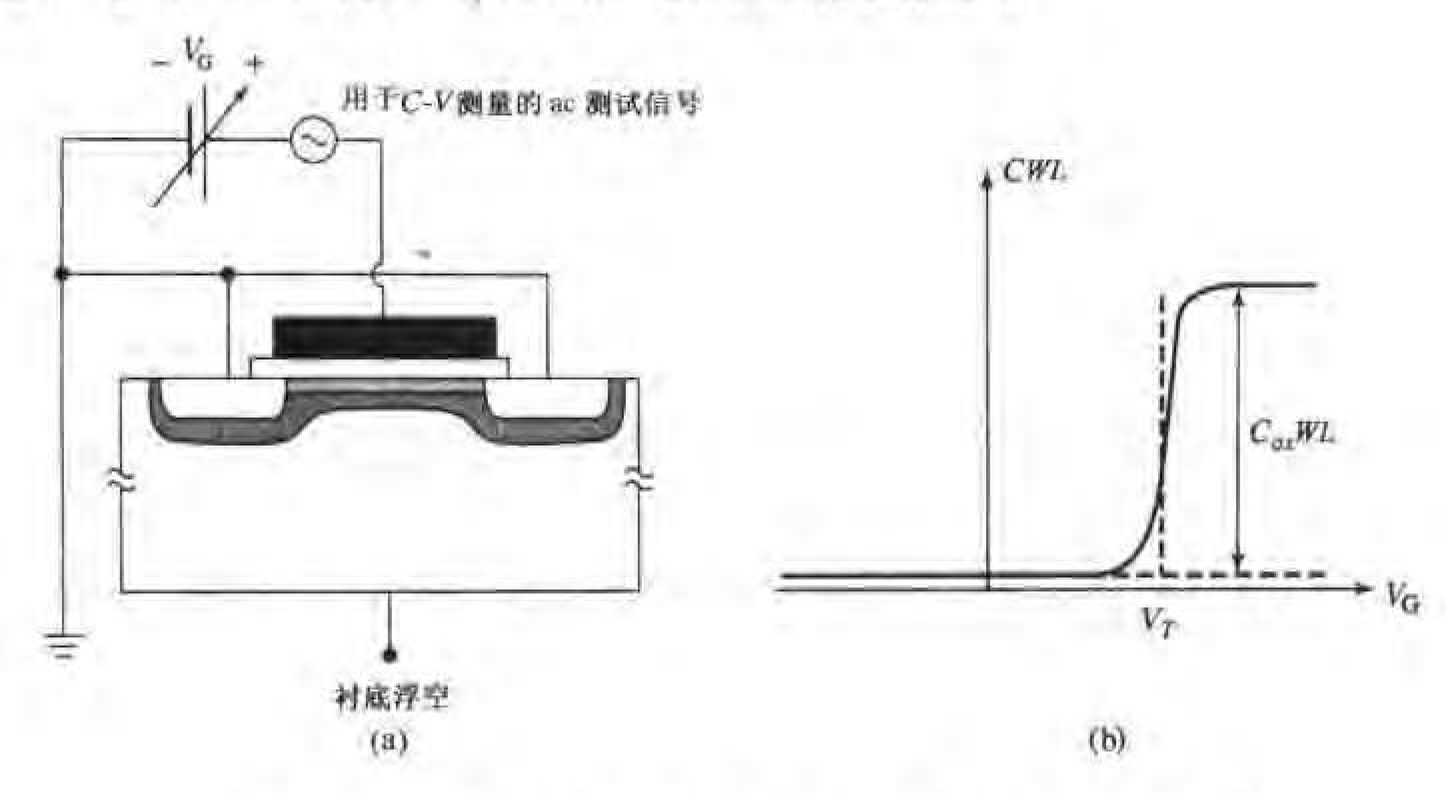


图 9.31 提取栅电容和栅氧化层厚度的(a)测量电路,(b)C-V 曲线

低场阈值电压 提取阈值电压 V_T的电路如图 9.32(a)所示,漏极电压固定在一个很小的值,通常约为 50mV,栅极电压 V_G不断增大,测量漏极电流。图 9.32(b)为三种不同衬偏电压 V_B下的测试曲线,曲线斜率最大处的切线与横轴的交点定义为阈值电压 V_T。这里的理论依据是 V_D[$V_D \ll 2(V_G - V_T)$]较小时等式(9.2.9)可以表示成

$$I_D \approx \mu_{eff} C_{ox} \frac{W}{L}(V_G - V_T)V_D \tag{9.3.2}$$

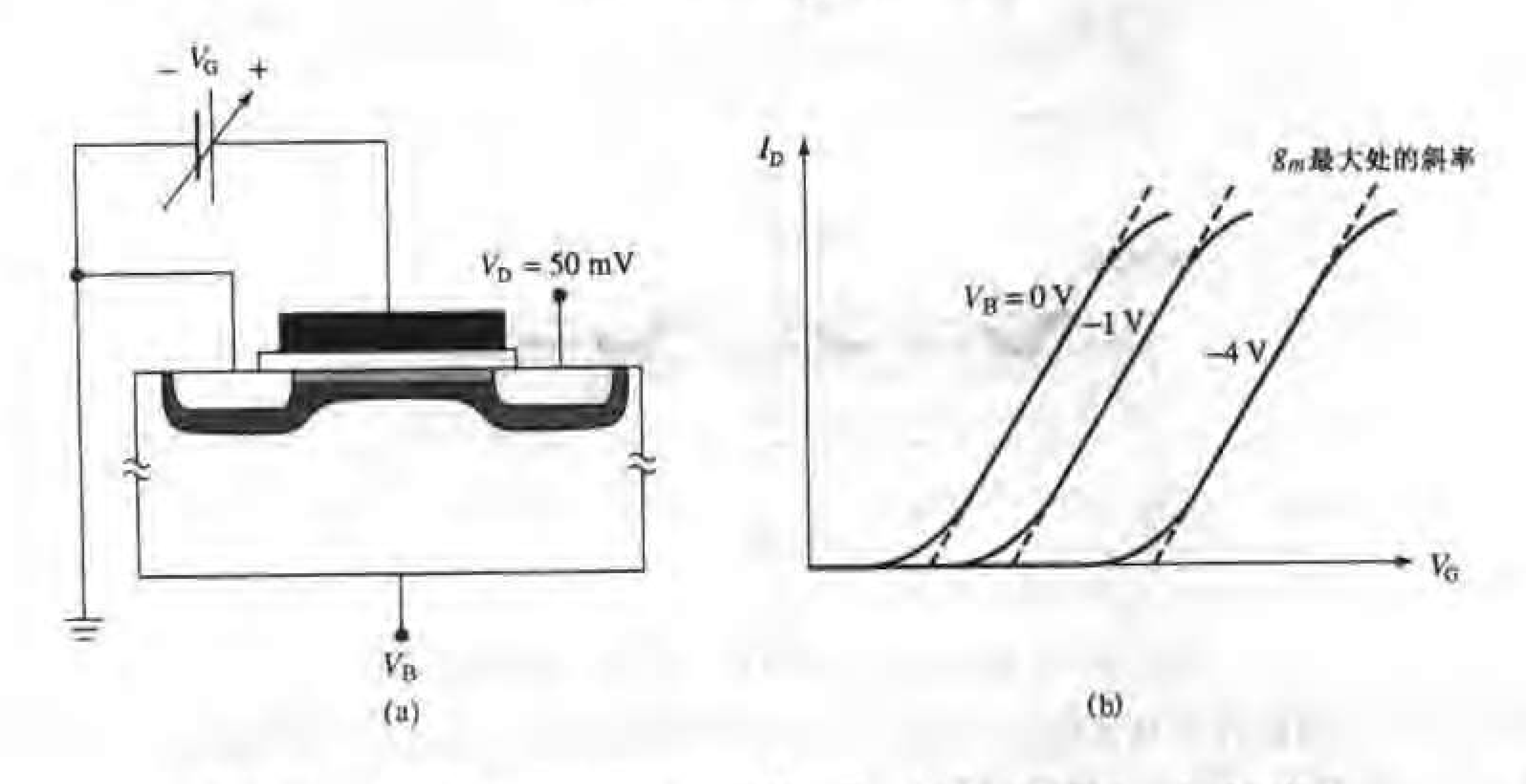


图 9.32 提取低场阈值电压的(a)测量电路,(b)$I_D - V_G$曲线。不同衬偏电压 V_B下的阈值电压,由直线段在 I_D = 0 处的截距求出

等式(9.3.2)表明,$V_G > V_T$时,I_D与 V_G呈线性关系;$V_G = V_T$时,I_D = 0(图 9.32(b)[③])。当栅压较

③ 译者注:原文误写为 9.31b。

高时，由于强电场的影响，载流子迁移率下降，电流曲线的斜率减小。

有效迁移率 图9.32(a)中的电路还可以用来提取有效迁移率。当漏极电压V_D很小时，栅-漏和栅-源电压几乎相同，因此从$y=0$到$y=L$的体电荷和沟道电荷都可以看做是常数。在这种情况下，迁移率也为常数。利用等式(9.3.2)求出迁移率的表达式为

$$\mu_{eff} \approx \frac{I_D L}{C_{ox} W (V_G - V_T) V_D} \tag{9.3.3}$$

要用这种方法提取精确的有效迁移率，必须知道等式(9.3.3)右端所有参数的准确值。这就带来一些问题，例如我们不知道MOSFET与沟道串联的电阻（源漏电极接触电阻和源漏扩散区电阻）。如果考虑串联电阻，那么沟道漏端电压将偏离外加漏极电压。与提高C_{ox}的提取精度一样，利用大器件尺寸可以减小与串联电阻有关的误差。

例题 参数提取

利用下表中n沟道MOSFET的测量数据提取阈值电压和迁移率。

	$V_G = 1W$	$V_G = 2V$
$V_D = 0.05V$	$I_D = 14\ \mu A$	$I_D = 34\ \mu A$

假设迁移率下降效应不明显，并且与沟道电阻相比，源漏电阻可以忽略。已知x_{ox} = 45nm，$W = 10\mu m$，$L = 1.0\mu m$，利用测量数据求MOSFET低场阈值电压和迁移率。

解：V_D很小时，等式(9.2.9)成立。利用表中的数据，有

$$14\ \mu A \approx \mu_{eff} C_{ox} \frac{W}{L}(1 - V_T)V_D \tag{1}$$

$$34\ \mu A \approx \mu_{eff} C_{ox} \frac{W}{L}(2 - V_T)V_D \tag{2}$$

等式(2)除以等式(1)，得到：$\frac{2-V_T}{1-V_T} \approx \frac{34}{14} \Rightarrow V_T = 0.3V$

求出阈值电压后，由等式(9.2.9)得

$$\begin{aligned} \mu_{eff} &\approx \frac{I_D L}{W C_{ox}(V_G - V_T)V_D} \\ &= \frac{14 \times 10^{-6} \times 1 \times 10^{-4}}{10 \times 10^{-4} \times \left(\frac{3.9 \times 8.85 \times 10^{-14}}{45 \times 10^{-7}}\right)(1 - 0.3)0.05} \\ &= 521\ cm^2V^{-1}s^{-1} \end{aligned}$$

9.3.3 CMOS的闩锁效应[+]

在设计CMOS电路时，必须避免闩锁(latch-up)效应，也就是由于寄生双极晶体管的作用，导致从电源到地出现低电阻箝位通路。为了分析该效应，考虑图9.33所示的p阱CMOS结构。从横截面示意图中可看出叠加在CMOS结构上的寄生的npn和pnp双极晶体管，这两个晶体管的集电结是共用的。图9.34为寄生双极晶体管的等效电路，在有效的偏压之下，pnp

晶体管的集电极提供 npn 晶体管的基极电流,同时 npn 晶体管的集电极提供 pnp 晶体管的基极电流。如果这两个晶体管具有中等的电流增益(βs),那么这种连接方式很容易使两个晶体管都进入饱和区,这样,电源和地之间由小电阻和两个电压降串联,这两个电压降是(1)基极-集电极饱和压降 V_{CEsat} 和(2)基极-发射极饱和压降 V_{BEsat}[①]。

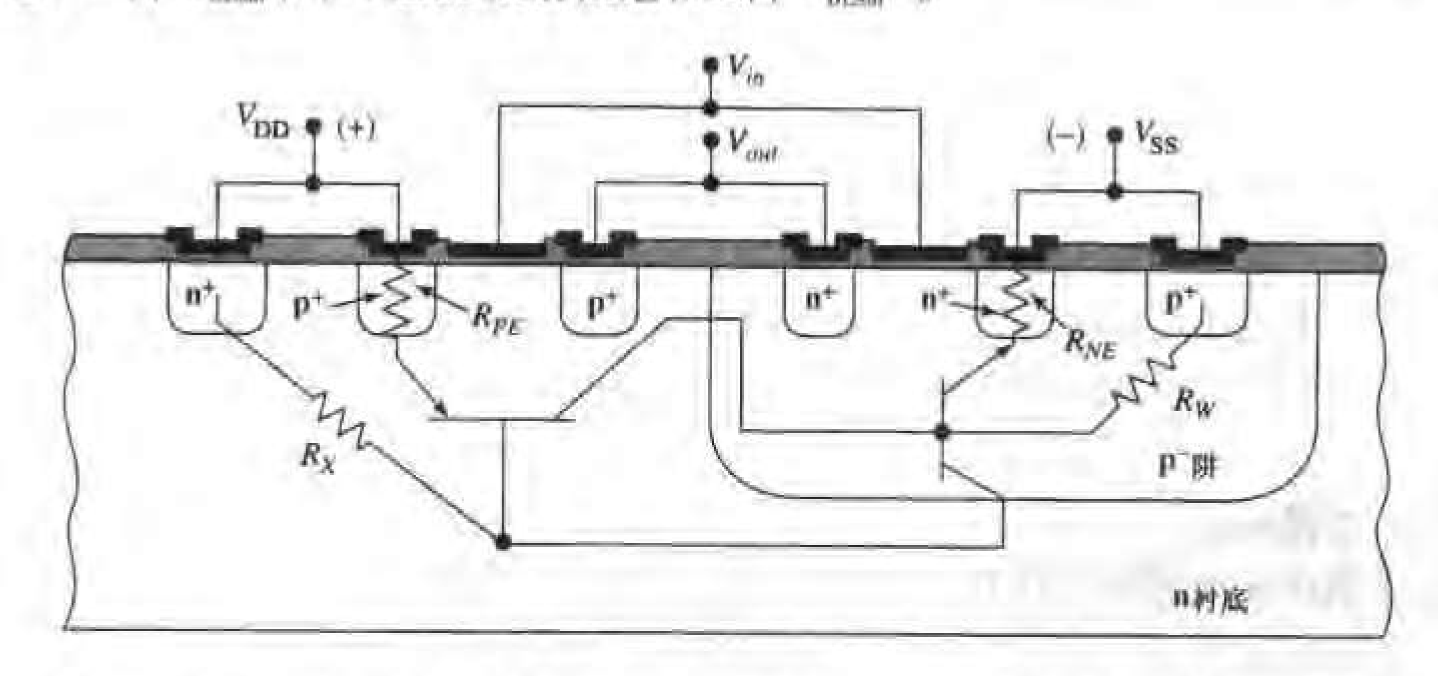


图 9.33 p 阱 CMOS 反相器横截面图,寄生的 pnp 和 npn 双极晶体管如图所示。R_X是衬底电阻,R_W是阱电阻,R_{PE}和 R_{NE}代表发射极处的电极接触和扩散区电阻

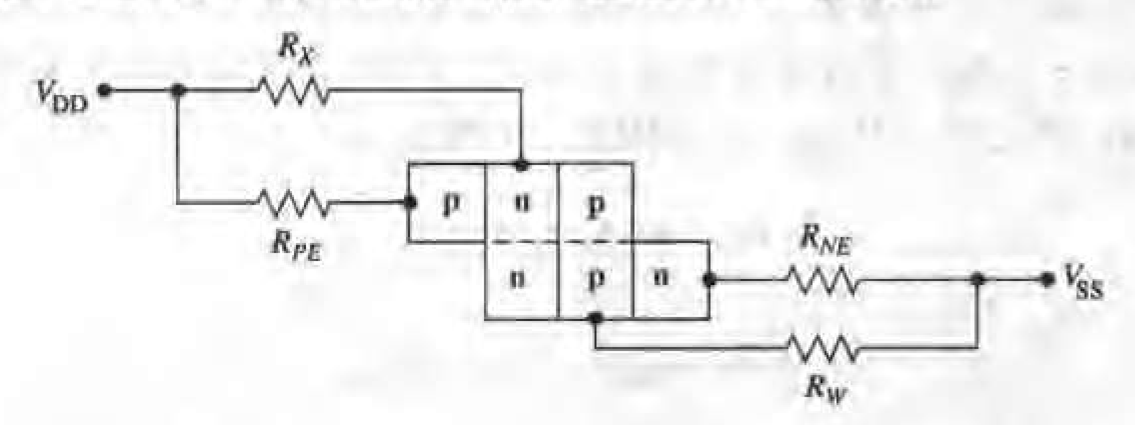


图 9.34 p 阱 CMOS 反相器中反映寄生 npn 和 pnp 晶体管效应的等效电路图

在通常的 CMOS 工作条件下,两个双极晶体管的发射结都处于反偏状态,不会出现闩锁效应。然而,一个好的电路设计必须保证在任何工作条件下都不会出现闩锁效应。首先我们来分析闩锁是如何产生的。参考图 9.35 所示的等效电路,该电路除了横向耦合的两个双极晶体管之外,还并联了电容 C_{PS}和电流源 I_0。电容 C_{PS}代表阱和衬底之间的 pn 结电容,它比典型的集电结电容大很多。电流源 I_0代表 pn 结的泄漏电流,通常情况下数值非常小。但是在某些机制下,I_0也会显著增大。

可能产生 I_0的原因有:(1)某些瞬态情况下,pn 结出现正偏,少数载流子注入衬底(通常发生在输入或输出电路中);(2)电离辐射形成的光致激发;(3)热载流子的碰撞电离。另外,大电容 C_{PS}在电压瞬变时也会产生电流,尤其是电路启动的过程中。上述任何一种电流,都有可

① 将 pnp 和 npn 晶体管连在一起,pnp 的基极由 npn 的集电极驱动,npn 的基极由 pnp 的集电极驱动,这是一种重要和有用的功率开关。这种开关通常称为可控硅整流器,简称 SCR。因为对 SCR 开关有较多的相关研究,通常 CMOS 中的闩锁效应也被称为 SCR 效应。

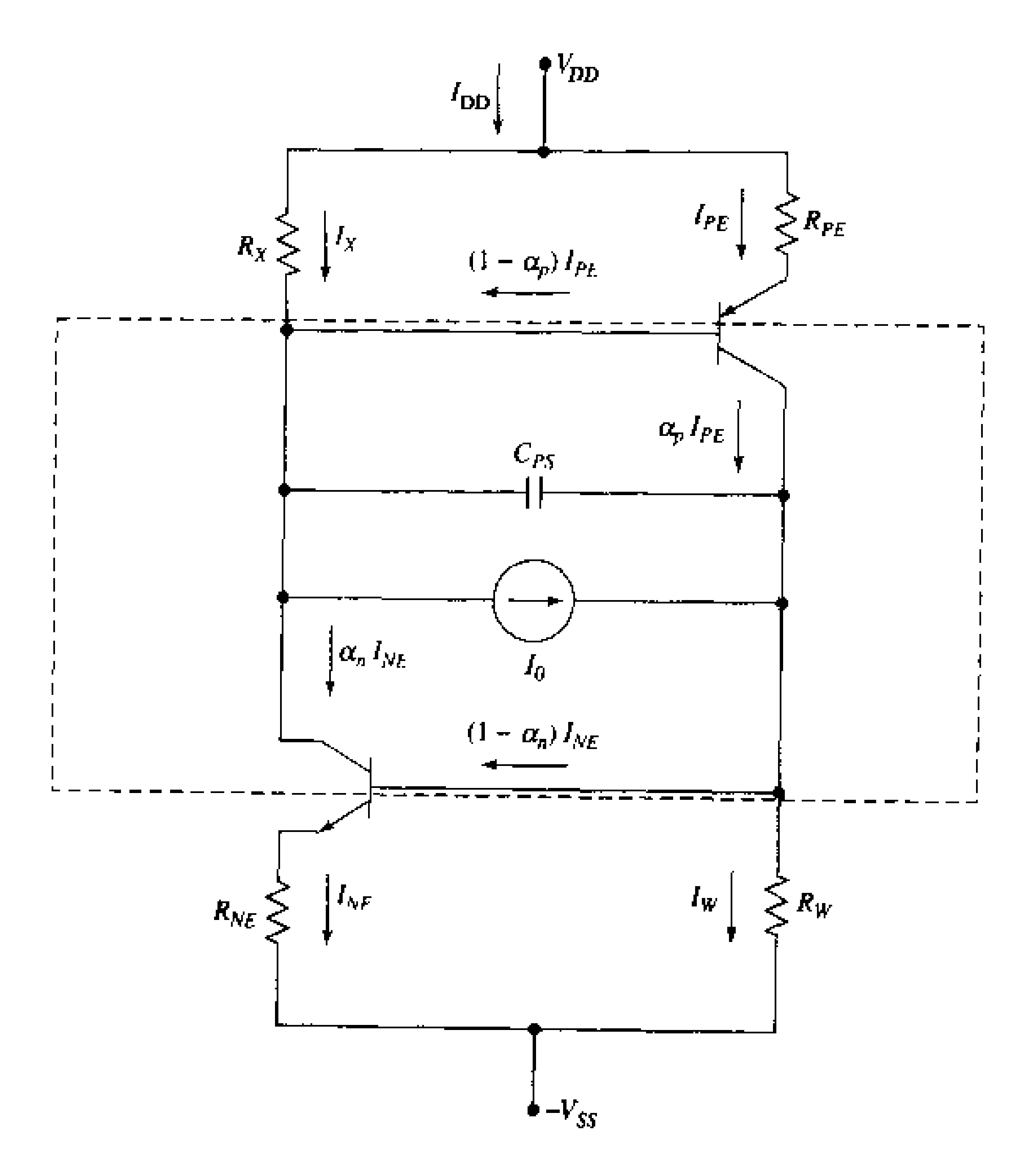


图9.35 闩锁效应等效电路,包括了阱-衬底电容 C_{PS} 和寄生电流源 I_0。虚线框里面包括了连接阱和衬底的所有元件

能导通一个或者两个寄生晶体管。如果寄生晶体管电流增益足够大,并且电源能够提供足够大的电流,闩锁效应就会发生。

闩锁效应模型 下面推导产生闩锁效应的条件。利用图9.35中的等效电路,忽略 I_0 和 C_{PS},并忽略 R_{PE} 和 R_{NE} 的作用(见图9.36)。在这样的条件下,驱动pnp晶体管基极的电流等于npn晶体管基极电流的 β_n 倍,被相互并联的pnp晶体管的输入电阻和衬底电阻 R_X 分流。在小信号条件下,pnp晶体管的基极电流与npn晶体管基极电流的比值为 $\beta_n R_X/(r_{\pi pnp}+R_X)$,其中 $r_{\pi pnp}$ 是 δg_m 的倒数(参见等式(7.5.3))。计算npn晶体管基极电流的方法与此类似,用阱电阻 R_W 代替 R_X[23]。这样,横向耦合的两个晶体管的环增益 G_L 为

$$G_L = \beta_n \times \frac{R_X}{r_{\pi pnp}+R_X} \times \beta_p \times \frac{R_W}{r_{\pi npn}+R_W} \tag{9.3.4}$$

在发生闩锁效应的情况下,环增益必须等于1;环增益小于1时则不会发生闩锁现象。简单的计算表明,要消除闩锁效应,就必须减小双极晶体管电流增益 βs,并尽可能减小 R_X 和 R_W 的值,这已成为避免闩锁效应的CMOS电路设计的基本指导方针。

在密度很高的CMOS电路中,为了更有效地避免闩锁效应,需要采用比较复杂的结构,而不是简单地在衬底上扩散一个阱形成"体CMOS"(见图9.29)。通常的做法是在高掺杂的衬

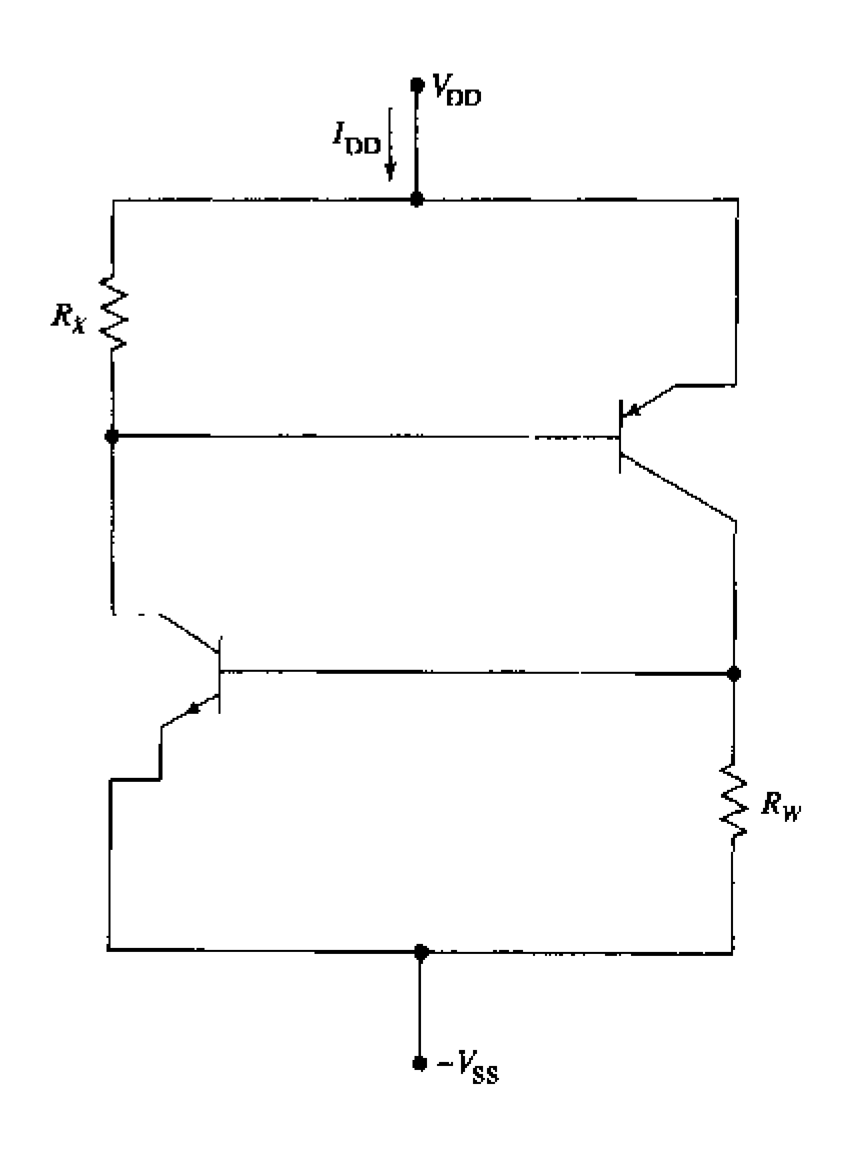


图 9.36 分析闩锁效应的简化模型

底上生长外延层，然后在外延层上做晶体管，这样就可以减小衬底电阻。同时，少数载流子的寿命在高掺杂的衬底中非常低，也能够降低寄生双极晶体管的增益。利用外延层结构可以减弱发生闩锁效应的可能性，即使在深亚微米电路中也是非常有效的。此外还可以通过辐射整个结构来降低少数载流子的寿命，从而降低双极晶体管的增益，但是这种方法同时会导致源漏结区载流子的寿命降低，增大了泄漏电流，这与 CMOS 的低功耗的优点是矛盾的。其他的措施还有重掺杂保护环。保护环可以在敏感位置起到箝位作用，或者使少数载流子在到达阱-衬底 pn 结之前就被收集。

例题 CMOS 中的闩锁效应

利用图 9.35 中的电路计算电源电流 I_{DD} 与阱电流 I_W、衬底电流 I_X 以及阱-衬底电流源 I_0 的关系。假设两个寄生晶体管都工作在放大状态，计算导致 I_{DD} 趋于无穷大的晶体管共基极电流增益 α。假设所有电压变化都足够慢。

解：因为电压变化足够慢，可以不考虑电容 C_{PS} 的作用。利用基尔霍夫电流定律，由图 9.35 得

$$I_{DD} = I_X + I_{PE}$$
$$I_{DD} = I_W + I_{NE}$$
$$I_W = \alpha_p I_{PE} - (1 - \alpha_n) I_{NE} + I_0$$
$$I_X = \alpha_n I_{NE} - (1 - \alpha_p) I_{PE} + I_0$$

消去 I_{PE} 和 I_{NE}，得

$$I_{DD} = I_X + \frac{(I_W - I_0)}{\alpha_p} + \frac{(1-\alpha_n)}{\alpha_p} \times \left[\frac{(I_X - I_0)}{\alpha_n} + \frac{(1-\alpha_p)}{\alpha_n} \times (I_{DD} - I_X)\right]$$

从上式求解 I_{DD},得

$$I_{DD} = \frac{I_0 - \alpha_p I_X - \alpha_n I_W}{1 - (\alpha_n + \alpha_p)}$$

从 I_{DD}的表达式我们发现,当$(\alpha_n + \alpha_p)$趋于 1 时,I_{DD}趋于无穷大(电路进入闩锁状态)。我们可以与等式(9.3.4)得出的结论相比较,等式(9.3.4)的闩锁约束条件是利用双极晶体管的小信号模型得到的,因此该结论是在瞬态建立过程中出现闩锁效应要求的电路增益。本例题考虑的是稳态情况,不需要考虑 r_π的分流,这样等式(9.3.4)可简化为$\beta_n \times \beta_p = 1$,或

$$\left(\frac{\alpha_n}{1-\alpha_n}\right)\left(\frac{\alpha_p}{1-\alpha_p}\right) = 1$$

同样可以得出结论$(\alpha_n + \alpha_p = 1)$。

9.4 展望未来

MOSFET 工艺中,表面最小尺寸是器件密度的特征尺寸。在超过 35 年的 MOSFET 发展历史中,这一尺寸,通常指的是 MOSFET 的沟道长度,在不断地缩小,有源器件的密度也在不断增大。正如我们在第二章中提到的,这一趋势通常用 Moore 定律描述。该定律在 20 世纪 70 年代早期由 Gordon Moore 提出,它预测 MOSFET 电路的集成度每 18~24 个月翻一番。直觉告诉我们,在这个快速发展的领域,最小尺寸最终是会到达极限的。20 世纪 80 年代人们预测 MOSFET 表面尺寸的最小值为 0.5μm,十年后这个预测变为 0.1μm,但是在今天,21 世纪的初期,人们常说的极限尺寸大约为 25nm。表 9.2 举例说明了 IC 工艺中的按比例缩小在过去和未来预测的发展趋势。

9.4.1 按比例缩小的目标

按比例缩小的主要目标是:(1)减小晶体管的尺寸;(2)提高单位宽度的驱动电流;(3)降低电源电压;(4)减小总负载电容。按比例缩小使我们可以在给定的硅片面积上集成更多的功能更强的电路。所有上述目标都可以提高性能,降低功耗,提供更廉价的产品。另外,按比例缩小必须考虑的约束条件是:(1)在关态时保证足够低的泄漏电流;(2)尽可能减小短沟效应;(3)保证电路的可靠性。评价按比例缩小的 MOSFET 性能的主要参数有:开态电流 I_{on},关态电流 I_{off},亚阈值斜率 S,输入阻抗 Z_{in},跨导 g_m和击穿电压 BV 等等。

9.4.2 栅耦合

减小栅氧化层厚度可以提高栅电压与沟道电荷之间的耦合,这正是我们所需要的。每一代新的 CMOS 技术都采用了更薄的栅氧化层。但是,栅氧化层的减薄也是有极限的。当栅氧

化层非常薄时,直接的电子隧穿电流会导致较大的栅电流,从而降低了输入阻抗。可被接受的栅氧化层厚度的极限大约是2.5nm。除了直接隧穿的限制外,栅氧化层厚度低于2.5nm并不能有效地提高栅极耦合作用,这是因为多晶硅栅电极中的载流子耗尽以及硅中的反型层(沟道)厚度有限,如图9.37所示。

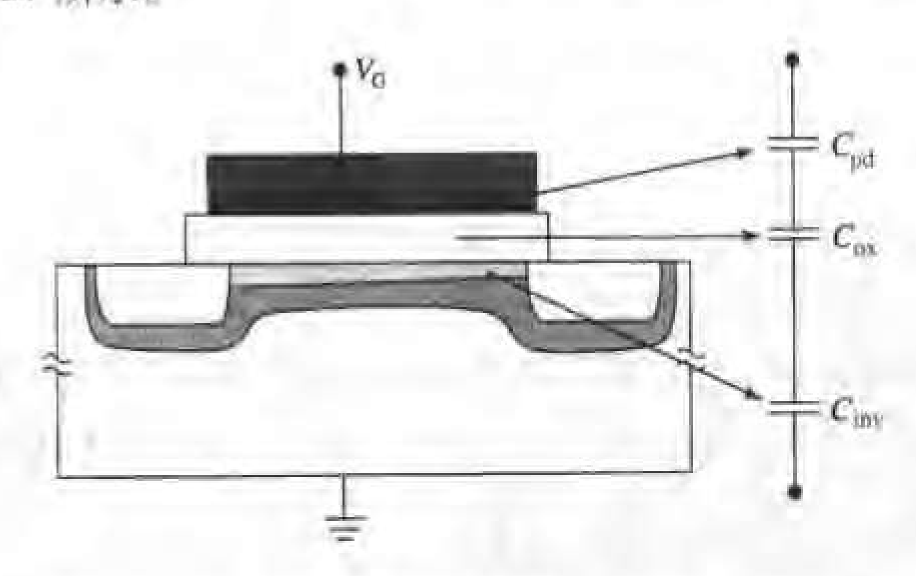


图9.37 MOSFET横截面示意图,栅电容C_G由三部分电容串联组成:氧化层电容C_{ox},多晶硅栅耗尽区电容C_{pd},有限厚度的导电沟道导致的电容C_{inv}

在推导MOSFET理论时,我们作了与栅-沟道耦合有关的几个假设:(1)假设多晶硅栅类似金属,不会出现载流子耗尽现象。当靠近氧化层一侧的部分多晶硅出现耗尽时,该耗尽层就类似绝缘层,增大了有效的栅绝缘层厚度,降低了栅电压对沟道的耦合作用。(2)假设反型层非常非常薄,并位于硅衬底的表面。但实际上反型层是有一定厚度的[30],降低了栅与沟道载流子之间的耦合作用,从而降低了栅的控制能力。当栅氧化层厚度较大时,上述假设是可行的。但是当栅氧化层很薄(例如,≤5nm)时,就会产生较大的误差。多晶硅耗尽电容和反型层电容的作用将使MOSFET的电流驱动能力大大下降。为了提高栅耦合,人们研究了一些新的技术,其中一个方法就是采用高介电常数(高κ)材料作为栅绝缘介质,栅电极材料采用金属硅化物。虽然这些新的技术还没有用于实际生产,但工程师们相信在未来的CMOS工艺中,高κ材料一定会发挥作用。

采用高κ介质的原因非常明显。单位面积栅氧电容的表达式为ϵ_{ox}/x_{ox},通过增大介电常数,而不是减小栅氧厚度来增大电容,可以避免直接隧穿发生。在过去,这种方法遇到的主要困难是不能保证淀积厚度的一致性,以及界面质量很差。然而,最近的研究表明,已经可以得到相当不错的界面。目前正在研究的高κ材料有:Al_2O_3($\epsilon_r=9.5$),ZrO_2($\epsilon_r=20\sim40$),HF_2O_2($\epsilon_r=20\sim30$),TiO_2($\epsilon_r=80$)以及Ta_2O_5($\epsilon_r=25$)。

采用高κ材料的同时,还可以采用金属或者金属硅化物作为栅电极材料,以减小多晶硅耗尽层的影响。对于n沟道MOSFET,要求金属的费米能级位置接近硅的导带(正如n^+多晶硅);对于p沟道MOSFET,则要求金属的费米能级在硅的价带附近(正如p^+多晶硅)。达到这一要求需要在CMOS电路中使用两种不同的金属,因此增加了工艺的复杂性。折衷的方法是采用费米能级靠近硅带隙中部的金属,但这样两种MOSFET都会有较大的阈值电压。通常我们采用的材料有双层W和TiN,以及三层Ta、TaN和NiSi。总之,金属栅工艺和CMOS多晶硅栅工艺并不兼容[24],如果要用金属栅替代多晶硅栅,则要求大量的工艺改进。

9.4.3 速度过冲

我们在9.2节讨论了速度饱和效应，其中假设载流子的平均自由程小于沟道长度，这样载流子可以将过剩的能量释放给晶格，达到饱和速度（第1章）。对于沟道非常短的器件，自由载流子速度可以大于饱和速度，这种现象称为速度过冲，将导致电流比我们的预期值大。

我们先回顾一下第1章的内容：自由载流子在电场中运动并获得能量，然后通过碰撞将能量传递给晶格。在低场下，载流子沿电场方向的平均速度随电场线性增加；在强场下，更有效的能量传输机制将起主要作用，有效地限制了平均速度，沿电场方向载流子的速度将达到饱和速度v_{sat}。假设两次碰撞之间载流子的平均自由程为λ，速度过冲现象发生在沟道长度与平均自由程相当或者更小的器件中。在这种情况下，通过碰撞损失能量的可能性降低，很多电子可以不损失能量地穿过整个沟道，因此载流子的平均速度就超过了v_{sat}，导致漏极电流大于预期值。

理论计算表明，速度过冲最终会变得非常重要。然而，由于载流子的平均自由程非常小，人们预测速度过冲只会发生在尺寸更小的器件中。我们在沟道非常短的晶体管中观察到的现象就可以用速度过冲加以解释[25~27]；但是，由于实验器件的尺寸和其他参数的不确定性，速度过冲假设目前很难得到验证。

趋势 CMOS工艺在集成电路中占主导地位，MOSFET也是使用最为广泛的现代集成电路器件。对于可以预见的未来，CMOS仍然是IC工艺的主流，器件尺寸也将不断缩小。在下一个十年，我们可以预测随着按比例缩小工艺的发展，器件的性能将进一步提高。但是我们还需要改进基本材料和工艺流程。

实验室已经可以做出沟道长度小于25nm的晶体管，但是，制作栅长小于100nm的MOSFET常常需要采用复杂的结构和不太熟悉的材料，这使得按比例缩小无法以早期的速度来提高电路的性能和密度。人们正在探索更为复杂的非常规的制作器件的方法，比如采用高介电常数（高κ）栅材料和金属栅电极。

小结

金属-氧化物-半导体场效应晶体管（MOSFET）的栅极电压可以控制源和漏之间的沟道电导，常被用做栅控电流开关，也可以用做放大器件，因为较低的栅源电压可以控制较高的源漏电流。栅源电压为零（$V_{GS}=0$）时就存在导电沟道的器件称为耗尽型MOSFET，栅极必须加上一定的电压才能形成导电沟道的器件称为增强型MOSFET。

MOSFET的漏极电流和漏极电压的关系（输出特性）可以分成两个区域：较低的漏极电压下，载流子在沟道中漂移运动，沟道从源到漏是连续的；较高的漏极电压下，导电沟道不会抵达漏端，在漏端附近，电流需要通过一段电场较强、载流子浓度较低的空间电荷区。一级近似表明，此时漏极电流达到饱和，MOSFET工作在饱和区。

简单的解析模型假设整个沟道都存在导电通路，因此该模型只在漏极电压小于饱和电压V_{Dsat}（使漏端附近的沟道自由电荷为零的漏极电压）时成立。当$V_D > V_{Dsat}$时，假设漏极电流不再受漏极电压的影响，而是决定于从源端到夹断点（导电沟道的终点）的电流。随着漏极电压的升高，夹断点向源端移动，导电沟道长度减小，称为沟道长度调制。该效应导致漏极电压大

于V_{Dsat}时,饱和漏极电流I_{Dsat}随着漏极电压的增大而升高。衬偏效应是指源和衬底之间的偏压改变时,器件的阈值电压将发生变化。当衬底掺杂浓度增大时,衬偏效应的影响加剧,但是随着氧化层电容的增大,该效应的影响将减小。

MOSFET 的跨导g_m随着漏极电压的升高而线性增大,直到器件饱和。一级近似中,饱和前跨导与栅极电压无关。饱和区的跨导g_m与栅极电压之间是线性关系,而与漏极电压无关。MOSFET 的速度通常不取决于沟道渡越时间,而是由器件中电容的充放电时间决定的。MOSFET 的等效电路由反映器件直流电流电压关系的元件以及电容和电阻组成。

MOSFET 的基本理论采用了很多近似处理,推导出的方程组给设计者带来了很多方便。最重要的简化是假设栅极电压小于阈值电压时,沟道中不存在有自由载流子。但是实际上,沟道中的自由载流子浓度在$V_G = V_T$这一点并不会发生突变,而是随栅极电压的减小而指数下降。$|V_G| < |V_T|$时,沟道中的自由载流子电荷导致 MOSFET 存在亚阈值电流。亚阈值电流的输运机制是扩散,而不是类似$|V_G| > |V_T|$时的漂移。

随着器件尺寸的减小或者电压的增大,基本 MOSFET 理论中的一些近似就不再成立,特别是自由载流子沿沟道的迁移率为常数的假设。虽然$|V_G| > |V_T|$时,漂移是主要的输运机制,但是迁移率沿沟道并不是常数,而是与沿沟道方向的电场$\mathscr{E}_x$和垂直于 $Si\text{-}SiO_2$界面的电场$\mathscr{E}_x$都有关系。即使在基本的长沟理论中,漏端附近的$\mathscr{E}_x$也会非常大,所以假设迁移率在整个沟道中是常数是有问题的。

短沟器件在较低的电压下就有可能出现速度饱和,此时大部分沟道的载流子漂移速度接近饱和速度。速度饱和导致漏极电流小于长沟理论预测值。由于速度饱和,I_{Dsat}随$(V_G - V_T)^2$的变化变缓,并且V_{Dsat}小于$(V_G - V_T)$。小尺寸 MOSFET 的阈值电压与沟道长度和沟道宽度都有关系。当沟道长度减小时,由于电荷共享,阈值电压减小。另外,漏致势垒降低和次表面穿通也会导致阈值电压减小。沟道宽度减小时,由于边缘电场的作用或掺杂离子的作用将导致阈值电压升高,而隔离结构造成的电力线的集中会导致阈值电压下降。

离子注入调整阈值电压是互补 MOS(CMOS)工艺中的关键步骤。CMOS 集成电路需要在同一块硅片上制作 n 沟道 MOSFET 和 p 沟道 MOSFET,这就要求在衬底上形成一个与衬底导电类型相反的阱。现在既有 p 阱工艺,也有 n 阱工艺,每一种工艺都有各自的优缺点。CMOS 数字反相器的直流功耗非常小,因为电路在稳定状态时从电源到地的泄漏电流非常小。CMOS 的这一特点在器件密度逐渐增大时显得尤为重要。但是对于高频工作的高密度的 IC,开关过程中的功耗是一个非常重要的问题。另外,CMOS 设计需要考虑的一个特殊问题是要避免闩锁效应,也就是寄生双极晶体管的正反馈作用使双极晶体管进入饱和区。

当器件尺寸缩小时,按比例缩小规律能够指导 MOSFET 的设计。其中一种按比例缩小规律是器件尺寸缩小时,保持器件中的电场强度不变,即恒场律。然而,实际情况限制了这一理论,因此,恒场律并没有按最初提出的形式使用。

正确预测器件按比例缩小的优缺点需要一个统一的、准确的、基于物理的短沟 MOSFET 模型。影响漏极电流模型的最重要的按比例缩小效应是氧化层电场和速度饱和造成的迁移率下降。如果采用较薄的栅氧化层厚度,可以增强栅电极与沟道载流子的耦合,使得晶体管特性更接近长沟器件。由于漏极电流和栅电容大约成正比,所以减薄栅氧化层厚度也有利于深亚微米工艺。短沟器件的截止频率f_T与沟道长度的倒数成正比,而长沟器件的f_T与沟道长度平方的倒数成正比。

参考文献

1. WILLIAM SHOCKLEY, *Electrons and Holes in Semiconductors*, D. Van Nostrand Co., Princeton, NJ, 1950, p. 29.
2. R. H. DENNARD, F. H. GAENSSLEN, H. N. YU, V. L. RIDEOUT, E. BASSOUS, and A. R. LEBLANC, *IEEE J. Solid-State Circuits*, **SC-9**, 256–68, 1974.
3. International Technology Roadmap for Semiconductors (1997–2001 editions), http://public.itrs.net
4. Z. H. LIU, C. HU, J. H. HUANG, T. Y. CHAN, M. C. JENG, P. K. KO, and Y. C. CHENG, *IEEE Trans. Electron Devices*, **ED-40**, 86–95, January 1993.
5. L. A. AKERS, M. M. E. BEGUWALA, and F. Z. CUSTODE, *IEEE Trans. Electron Devices*, **ED-28**, 1490–1495, Dec. 1981.
6. J. Y CHEN and R. C. HENDERSON, *IEEE IEDM Technical Digest*, 1982, p. 233–36.
7. L. A. AKERS, *IEEE Electron Dev. Lett.*, **EDL-7**, 419–21, July 1986.
8. K. OHE, S. ODANAKA, K. MARIYAMA, T. HORI, and G. FUSE, *IEEE Trans. Electron Devices*, **ED-36**, 1110–1116, June 1989.
9. R. R. TROUTMAN, *IEEE J. Solid-State Circuits*, **SC-14**, 383–91, 1979.
10. Y. C. CHENG and E. A. SULLIVAN, *J. Appl. Phys.* **45**, 187, 1974.
11. A. G. SABNIS and J. T. CLEMENS, *Tech. Dig.*, Int Electron Devices Meeting, 1979, pp. 18–21.
12. S. C. SUN and J. D. PLUMMER, *IEEE Trans. Electron Devices*, **ED-27**, 1497–1508, 1980.
13. M. S. LIANG, J. Y. CHOI, P. K. KO, and C. HU, *IEEE Trans. Electron Devices*, **ED-33**, 409, 1986.
14. K. K. HUNG, *Electrical characterization of the Si–SiO₂ interface for thin oxides*, Doctoral thesis, Dept of Electrical Engineering, Hong Kong University, 1987.
15. F. FANG and X. FOWLER, *J. Appl. Phys.* **41**, 1825, 1969.
16. R. COEN and R. S. MULLER, *Solid-State Electron.* **23**, 35–40, 1980.
17. J. A. COOPER and D. F. NELSON, *IEEE Electron Device Lett.* **EDL-2**, 169–73, 1983.
18. C. G. SODINI, P. K. KO, and J. L. MOLL, *IEEE Trans. Electron Devices*, **ED-31**, 1386, 1984.
19. M. C. JENG, J. CHUNG, A. WU, J. MOON, T. Y. CHAN, G. MAY, P. K. KO, and C. HU, *IEEE IEDM Tech. Digest*, 1987, p. 710.
20. F M. WANLASS and C. T. SAH, *IEEE Int. Solid-State Circuits Conf.*, Philadelphia, PA (Feb. 1963).
21. B. SHEU, D. L. SCHARFETTER, P. K. KO, and M. C. JENG, *IEEE Journal of Solid-State Circuits*, **SC-22**, p. 558–66, April 1987
22. The Compact Modelling Council: http://www.eigroup.org/cmc
23. K. W. TERRILL, *CMOS Latch-up Modeling and Prevention*, Doctoral thesis, Department of EECS, University of California, Berkeley, Dec. 1985.

习题

9.1 已知n沟道MOSFET和p沟道MOSFET的衬底掺杂浓度N_a和N_d分别为10^{15},10^{16},10^{17}和$10^{18}cm^{-3}$,制表说明阈值电压(考虑$V_S=V_B=0$)与掺杂浓度的关系。假设所有情况下Si-SiO$_2$界面固定电荷密度为$Q_f/q=10^{11}cm^{-2}$,栅氧化层厚度为20nm,栅电极为n$^+$多晶硅,其费米能级位于导带底。在表中指出MOSFET是增强型还是耗尽型。

9.2 假设MOSFET处于夹断状态,夹断区开始于沟道电压为V_{Dsat}的点,终止于漏区。假设夹断区完全耗尽,漏区掺杂浓度远高于沟道掺杂,推导等式(9.1.9)。

9.3* 已知n沟道MOSFET的$W/L=5$,栅氧化层厚度为20nm,迁移率为常数,其值为$\mu_n=600cm^2V^{-1}s^{-1}$,该器件用做可控电阻

(a)如果要求低V_{DS}时,MOSFET源漏之间的电阻为500Ω,计算沟道中的自由电子浓度Q_n/q。

(b)利用长沟理论计算达到(a)中要求的电阻所需要的栅电压与阈值电压的差(V_G-V_T)。

9.4* 已知n沟道MOSFET的$L=0.8\mu m$,$x_{ox}=15nm$,$V_T=0.7V$,$V_{GS}-V_T=3V$,$V_{DS}=2V$。

(a)利用长沟理论计算源端和漏端的沟道载流子速度,取$\mu_n=670cm^2V^{-1}s^{-1}$。

(b)利用9.2节更精确的模型重复(a)。取$v_{sat}=8\times10^6cm\ s^{-1}$。

9.5* 考虑耗尽型n沟道MOSFET的衬偏效应,电路连接方式与9.1节的例题相同。计算$V_{SB}=2V$、3V、4V、4.5V时的I_D。(注意,在这些条件下MOSFET没有饱和)

9.6 已知 $V_{SS}=6V$、8V、10V，重复9.5题。(注意，在这些条件下源漏互换，并且 $V_{GD}=0$)

9.7* n沟道MOSFET的一系列测试数据见下表：

V_{GS}(V)	V_{DS}(V)	V_{SB}(V)	I_D(μA)
3	4	0	120
3	6	0	130
3	4	4	76.8
4	4	0	270

利用等式(9.1.10)和(9.1.11)确定MOSFET的参数 $V_T(0)$、$k=\mu_n C_{ox} W/L$、γ 和 V_A。假设 $2|\phi_p|=0.6V$(注意，ϕ_p 与 N_a 的依赖关系较弱)。

9.8 (a)已知 n 沟道MOSFET的栅氧化层厚度为50nm，衬底掺杂浓度 $N_a=2\times10^{15}cm^{-3}$，平带电压 $V_{FB}=-0.2V$，$\mu_n W/L=5\times10^3 cm^2V^{-1}s^{-1}$，比较等式(9.1.6)和(9.1.16)计算出的 I_{Dsat}。考虑 $V_{GS}=5.5V$、4.5V、3.5V和2.5V。

(b)利用(a)中用等式(9.1.16)计算出的 I_{Dsat}，确定 $V_{GS}=3.5V$ 时等式(9.1.19)中的参数 α，使两个等式计算出的 I_{Dsat} 相等。利用计算出来的 α，比较(a)中的其他 V_{GS} 下等式(9.1.16)和(9.1.19)计算出的 I_{Dsat}。

9.9 9.1节讨论了离子注入调整阈值电压。利用"方盒"近似推导等式(9.1.21)。

9.10 对于离子注入MOSFET，按文中指出的步骤推导反型时的耗尽层电荷浓度等式(9.1.22)和阈值电压等式(9.1.23)。

9.11 考虑偏置在强反型线性区的MOSFET。

(a)写出以 y 为变量的沿沟道方向的扩散电流表达式。

(b)写出以 y 为变量的沿沟道方向的漂移电流表达式

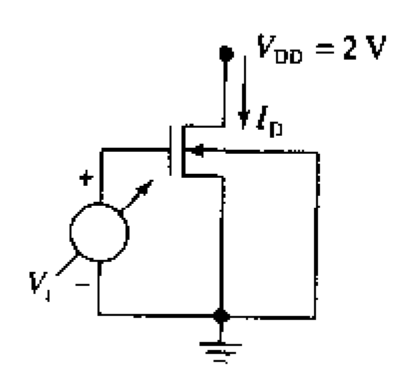
(c)比较(a)和(b)中的表达式，请问哪个电流成分起主要作用？[提示：需要简化，并利用Einstein关系 $D_n/\mu_n=kT/q$]

9.12 考虑图P9.12所示的电路，MOSFET由等式(9.1.5)和(9.1.6)描述，已知 $\mu_n C_{ox} W/L=25\mu A/V^2$，$V_T=1V$，$W/L=2$，电压 V_i 在0V和4V之间变化。

(a)绘出 $\sqrt{I_D}\sim V_i$ 关系曲线，并说明曲线中的突变点。

(b)用实线绘出MOSFET的跨导曲线。

(c)在(b)问的图中，用虚线绘出输出电导曲线($g_D=\partial I_D/\partial V_{DS}$)。



图P9.12

9.13 利用等式(9.1.3)推导出饱和区 Q_n 的表达式，并利用该式证明饱和区栅源小信号电容为

$$C_{GS}=\left|\frac{\partial Q_n}{\partial V_{GS}}\right|=\frac{2}{3}C_{ox}WL$$

9.14 证明表9.1中的恒场律按比例缩小规律。

9.15 证明对于长沟厚栅氧MOSFET，9.2节推导的精确 I-V 模型(等式(9.2.9)、(9.2.10)、(9.2.11)和(9.2.14))收敛于经典长沟模型(等式(9.1.5)、(9.1.6)和(9.1.31))。

9.16* 公司 A 开发的亚微米 CMOS 工艺的规格为：$x_{ox}=15\text{nm}$，$x_j=0.2\mu\text{m}$，n 沟道 MOSFET 为 n^+ 多晶硅栅，$V_T=0.7\text{V}$；p 沟道 MOSFET 为 p^+ 多晶硅栅，$V_T=-0.7\text{V}$。如果器件的 $L_{eff}=0.5\mu\text{m}$，$W=100\mu\text{m}$，在 V_{DD} 分别为 3.3V 和 5.0V 时估算 I_{Dsat} 和 $g_{msat}(=dI_{Dsat}/dV_G)$。如果器件的 $L_{eff}=0.02\mu\text{m}$，I_{Dsat} 可以提高多少百分比（不考虑穿通）？

9.17 已知器件参数为：$x_{ox}=20\text{nm}$，$L=1.0\mu\text{m}$，$v_{sat}(\text{old})=10^7\text{cm/s}$，$\mu_{eff}(\text{old})=500\text{cm}^2/\text{V-s}$，$V_T=0\text{V}$。对比以下两种提高 I_{Dsat} 的方法：

(a) 饱和速度 v_{sat} 提高为原来的 2 倍

(b) 有效迁移率 μ_{eff} 提高为原来的 2 倍（v_{sat} 保持不变）

通过计算绘出这两种情况下的 $I_{Dsat}(\text{new})/I_{Dsat}(\text{old})$ 与 V_G 的关系曲线。假设 μ_{eff} 与 V_G 无关

9.18 设 MOSFET 的 x 方向垂直于表面并从表面指向硅，y 方向从源指向漏，$V_D<V_{Dsat}$ 时 $Q_n(y)=[V_G-V_T-V(y)]$。

(a) 写出 $x=0$ 处的 $V(y)$ 以及横向电场 $\mathscr{E}_y(y)$ 的表达式；

(b) 用 (a) 中推导出的表达式求 MOSFET 处于临界饱和状态时的 $V(y)$ 和 $\mathscr{E}_y(y)$。

(c) 利用 (a) 和 (b) 的结论（不考虑是否正确），绘出 $V_G>V_T$，$V_D=0.5V_{Dsat}$ 时的 $\mathscr{E}_y(y)$ 与 y 的关系曲线，y 的范围从 0 到 L。

9.19 考虑图 9.28(a) 中的 CMOS 反相器。

(a) 在图 9.28(c) 中的电压传输特性（VTC）曲线上指出 v_I 改变时各 MOSFET 的状态，例如，v_I 为零时，p 沟道 MOSFET 为欧姆电阻，n 沟道 MOSFET 截止。在 VTC 曲线上指出 MOSFET 导通状态改变的所有点。

(b) 如果两个 MOSFET 的特性都可以用等式(9.1.5)和(9.1.6)描述，计算(a)中指出的所有点的电压。已知 n 沟道 MOSFET 的 $\mu_n C_{ox} W/L=40\mu\text{A/V}^2$，$V_T=1\text{V}$；p 沟道 MOSFET 的 $\mu_p C_{ox} W/L=35\mu\text{A/V}^2$，$V_T=-1\text{V}$。两个晶体管的 $\gamma=1/V_A=0$，电源电压 $V_{DD}=5\text{V}$。

9.20 已知图 P9.20 所示的增强负载反相器电路，两个晶体管的特性都可以用等式(9.1.5)和(9.1.6)描述，$\mu_n C_{ox} W/L=40\mu\text{A/V}^2$，$V_T=2\text{V}$，电源电压 $V_{DD}=8\text{V}$，忽略衬偏效应和沟道长度调制（也就是说，$\gamma=1/V_A=0$）。注意，输出电压 v_o 不会超过 $(V_{DD}-V_T)$，因为 v_o 高于这个数值时，不会有电流流过负载。

(a) 下方的 MOSFET 的栅极电压 v_I 分别为 0V，2V，4V 和 6V，漏极电压在 0～8V 之间变化，绘出其输出特性曲线（$I_D\sim V_{DS}$）。

(b) 在(a)中绘出的曲线上画出电路的负载线，也就是输出电压 v_o（下方的 MOSFET 的 V_{DS}）与输入电压 v_I 之间的关系。

(c) 如果负载改用 $-20\text{k}\Omega$ 的电阻，重复(b)

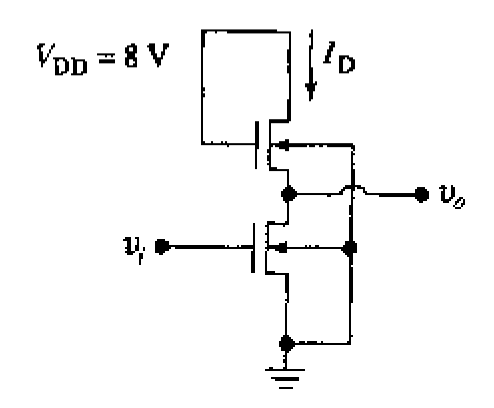


图 P9.20

9.21* 考虑图 P9.21 所示的耗尽负载反相器，两个晶体管的特性都可以用等式(9.1.5)和(9.1.6)描述，增强型 MOSFET 的 $\mu_n C_{ox} W/L=50\mu\text{A/V}^2$，$V_T=1\text{V}$；耗尽型 MOSFET 的 $\mu_n C_{ox} W/L=10\mu\text{A/V}^2$，$V_T=-3\text{V}$。与图 P9.20 相比，该电路的输出电压可以达到电源电压 $V_{DD}=5\text{V}$，因为 $V_{GS}=0\text{V}$ 时负载导电。

(a) 不考虑衬偏效应和沟道长度调制，如果 $v_I=5\text{V}$，计算 v_o。

(b) 假设体效应参数 $\gamma=0.4\text{V}^{1/2}$，$|\phi_p|=0.3\text{V}$，计算 v_o 达到其最大值时负载器件的阈值电压。

(c)利用(b)中的 γ 重复(a)中的计算.

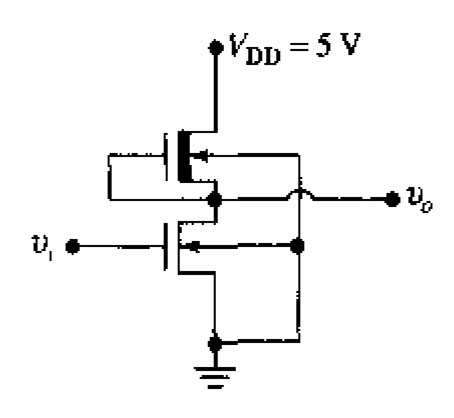


图 P9.21

9.22' 利用图 9.36 证明为保证 CMOS 电路处于闩锁状态,电源 V_{DD} 必须提供的电流为

$$I_{DD}=\frac{(V_{BE}/R_W)\beta_n(\beta_p+1)+(V_{BE}/R_X)\beta_p(\beta_n+1)}{(\beta_n\beta_p-1)}$$

(本题表明可以通过限制电源提供的电流来避免 CMOS 电路的闩锁)

9.23 考虑 9.2 节例题中描述电荷共享的简单几何模型.推导减小因子等式(2)和阈值电压等式(3).

9.24' (a)在 MOSFET 的沟道区考虑 Poisson 方程(等式(4.1.10)),假设衬底保持电中性,证明杂质浓度项(N_d-N_a)可以写成 $n_i[\exp u_B-\exp(-u_B)]=2n_i\sinh u_B$,其中 u 指的是用(kT/q)归一化的电势,u_B 为 p 型衬底的电势($u_B=-q\phi_b/kT$).

(b) 利用(a)中的表达式,证明 Poisson 方程可以用归一化电势写成

$$\frac{d^2u}{dx^2}=\frac{1}{L_{Di}^2}[\sinh(u)-\sinh(u_B)]$$

式中 L_{Di} 为 Debye 长度

$$L_{Di}=\left(\frac{\epsilon_s kT}{2q^2n_i}\right)^{1/2}$$

(c)计算 300K 时的 Debye 长度 L_{Di}

9.25' (a)求解 9.24 题推导出来的 Poisson 方程.将等式两边同乘积分因子 $2du/dx$,这样等式的左边可以化成全微分项

$$\left[\frac{du}{dx}\right]^2=\mathscr{E}^2$$

(b)证明 Si-SiO$_2$ 界面 $u=u_s$ 处硅中的电场 $\mathscr{E}_s$ 为

$$\mathscr{E}_s=\pm\frac{kT}{qL_{Di}}F_s(u_s,u_B)$$

式中(u_B-u_s)>0 时取正号,(u_B-u_s)<0 时取负号,函数 F_s 的定义为

$$F_s\equiv\sqrt{2}[(u_B-u_s)\sinh u_B-(\cosh u_B-\cosh u_s)]^{1/2}$$

(c)利用能带图证明(u_B-u_s)>0 对应 p 型衬底的积累情形,(u_B-u_s)<0 对应耗尽和反型情形

9.26' 硅表面电荷 Q_s 与 $\mathscr{E}_s$ 的关系为 $Q_s=\epsilon_s\mathscr{E}_s$,因此

$$Q_s=\pm\frac{\epsilon_s kT}{qL_{Di}}F_s$$

式中正、负号的确定与 9.25 题相同.已知衬底掺杂浓度 $N_a=10^{15}\text{cm}^{-3}$,$T=300\text{K}$,在半对数坐标系中绘出

$|Q_s|$ 与表面势 u_s 的关系曲线，考虑表面积累（$u_s - u_B$ 为正）、平带（$u_s = u_B$）和反型（$u_s > -u_B$），u_s 的范围从 −20 到 +20，在图中标出平带电压和阈值电压。［9.24 题和 9.26 题提供了利用 Poisson 方程表示表面电荷、严格求解亚阈值电流的基础］

表 9.4 MOSFET 公式

n⁻沟道	p⁻沟道
MOSFET 公式	
开启时的耗尽电荷密度（等式(8.3.9)）	
$Q_d = -qN_a x_{dmax} = -\sqrt{2\epsilon_s qN_a(2\|\phi_p\| + \|V_{SB}\|)}$	$Q_d = +qN_d x_{dmax} = +\sqrt{2\epsilon_s qN_d(2\|\phi_n\| + \|V_{SB}\|)}$
平带电压（等式(8.5.6)）	
$V_{FB} = \Phi_{MS} - \frac{Q_I}{C_{ox}} - \frac{1}{C_{ox}}\int_0^{x_{ox}} \frac{x\rho(x)}{x_{ox}}dx$	
阈值电压（等式(8.3.18)）	
$V_T = V_{FB} + V_S + 2\|\phi_p\| + \frac{\|Q_d\|}{C_{ox}}$	$V_T = V_{FB} - V_S - 2\|\phi_n\| - \frac{\|Q_d\|}{C_{ox}}$
考虑衬偏的阈值电压漂移（等式(9.1.11)）	
$\Delta V_T = \frac{\sqrt{2\epsilon_s qN_a}}{C_{ox}}\left(\sqrt{2\|\phi_p\| + \|V_{SB}\|} - \sqrt{2\|\phi_p\|}\right)$	$\Delta V_T = -\frac{\sqrt{2\epsilon_s qN_d}}{C_{ox}}\left(\sqrt{2\|\phi_n\| + \|V_{SB}\|} - \sqrt{2\|\phi_n\|}\right)$
长沟电流电压公式	
线性区漏极电流（等式(9.1.5)）	
$I_D = \mu_n C_{ox}\frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}\right]$	$I_D = -\mu_p C_{ox}\frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}\right]$
长沟饱和漏极电压	
$V_{Dsat} = (V_{GS} - V_T)$	
饱和区漏极电流（等式(9.1.6)）	
$I_{Dsat} = \mu_n C_{ox}\frac{W}{2L}(V_{GS} - V_T)^2$	$I_{Dsat} = -\mu_p C_{ox}\frac{W}{2L}(V_{GS} - V_T)^2$
沟道长度调制（等式(9.1.10)）	
$I_{Dsat} = \mu_n C_{ox}\frac{W}{2L}(V_{GS} - V_T)^2\left(1 + \frac{V_{DS}}{V_A}\right)$	$I_{Dsat} = -\mu_p C_{ox}\frac{W}{2L}(V_{GS} - V_T)^2\left(1 - \frac{V_{DS}}{V_A}\right)$
饱和区跨导（等式(9.1.37)）	
$g_{msat} = \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T) = \frac{2I_{Dsat}}{(V_{GS} - V_T)}$	$g_{msat} = -\mu_p C_{ox}\frac{W}{L}(V_{GS} - V_T) = -\frac{2I_{Dsat}}{(V_{GS} - V_T)}$
考虑体电荷的长沟电流电压公式	
线性区漏极电流（等式(9.1.17)）	
$I_D = \mu_n C_{ox}\frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \frac{\alpha V_{DS}^2}{2}\right]$	$I_D = -\mu_p C_{ox}\frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \frac{\alpha V_{DS}^2}{2}\right]$
饱和漏极电压（等式(9.1.18)）	
$V_{Dsat} = \frac{(V_{GS} - V_T)}{\alpha}$	

续表

MOSFET 公式	
n^- 沟道	p^- 沟道
饱和漏极电流(等式(9.1.19)) $I_{Dsat}=\mu_n C_{ox}\frac{W}{2\alpha L}(V_{GS}-V_T)^2$	$I_{Dsat}=-\mu_p C_{ox}\frac{W}{2\alpha L}(V_{GS}-V_T)^2$
短沟电流电压公式	
有效纵向电场(等式(9.2.3)) $\mathscr{E}_{eff}=\frac{(V_{GS}-V_T)}{6x_{ox}}+\frac{(V_T+V_Z)}{3x_{ox}}$	$\mathscr{E}_{eff}=\frac{(V_{GS}-V_T)}{6x_{ox}}+\frac{(V_T-V_Z)}{3x_{ox}}$
有效迁移率(等式(9.2.4)) $\mu_{eff}=\frac{\mu_0}{1+(\mathscr{E}_{eff}/\mathscr{E}_0)^\nu}$	$\mu_{eff}=\frac{\mu_0}{1+(-\mathscr{E}_{eff}/\mathscr{E}_0)^\nu}$
饱和电场强度(等式(9.2.7)) $\mathscr{E}_{sat}=2v_{sat}/\mu_{eff}$	$\mathscr{E}_{sat}=-2v_{sat}/\mu_{eff}$
线性区漏极流(等式(9.2.9)) $I_D=\mu_{eff}C_{ox}\frac{W}{L}\left[(V_{GS}-V_T)-\frac{V_{DS}}{2}\right]\frac{V_{DS}}{1+(V_{DS}/\mathscr{E}_{sat}L)}$	$I_D=-\mu_{eff}C_{ox}\frac{W}{L}\left[(V_{GS}-V_T)-\frac{V_{DS}}{2}\right]\frac{V_{DS}}{1+(V_{DS}/\mathscr{E}_{sat}L)}$
饱和漏极电压(等式(9.2.11)) $V_{Dsat}=\frac{(V_{GS}-V_T)\mathscr{E}_{sat}L}{V_{GS}-V_T+\mathscr{E}_{sat}L}$	
饱和区漏极电流(等式(9.2.10)) $I_{Dsat}=WC_{ox}v_{sat}(V_{GS}-V_T-V_{Dsat})$	$I_{Dsat}=-WC_{ox}v_{sat}(V_{GS}-V_T-V_{Dsat})$

第 10 章　MOS 场效应晶体管 II:强场效应

第 9 章描述了 MOSFET 的基本工作原理,首先利用缓变沟道近似分析了长沟器件,并推导出正常偏置下的电流电压方程,然后介绍了由速度饱和与迁移率下降导致的效应以及短沟、窄沟效应。尽管第 9 章中的理论可以描述多种情况下 MOSFET 的行为,但是对于一些发生在速度饱和区的重要的物理效应,上述理论中的近似限制了其准确性,尤其是没有考虑强场效应。强场效应会限制常规 MOSFET 的性能,但在某些特殊器件中却可以利用该效应。

本章将研究当今 MOSFET 中的强电场如何导致载流子进入正常的导电沟道以外的区域,将集中分析以下效应:(1)沟道长度调制导致的输出电阻减小;(2)从沟道流入衬底电极的电流;(3)从沟道经过栅氧化层流到栅电极的电流。本章还将讨论如何避免这些效应。MOSFET 的按比例缩小通常会使器件中出现明显的强电场,其峰值在漏区附近。沟道中的载流子在强电场中能获得较高的能量,可能在漏区附近的表面区域导致 MOSFET 不可逆转的性能退化,这种随时间积累的退化最终将导致器件失效。

因为这一系列对 MOSFET 的按比例缩小和器件性能的限制取决于器件中的电场,尤其是峰值电场的大小,所以首先分析电场,求出比第 9 章的讨论更为精确的电场分布。我们将利用"准二维"方法求解速度饱和区的 Poisson 方程,并分析相关的物理机制,然后通过对 MOSFET 电场的分析来考虑衬底电流和栅极电流。接着将讨论这些不需要的电流可能导致的器件退化,并比较 n 沟道和 p 沟道 MOSFET 物理效应的不同。本章的最后讨论强电场将电荷注入氧化层的用途,即用于制造在断电情况下仍能保留信息的不挥发 MOS 存储器。

10.1　速度饱和区的电场

当短沟 MOSFET 的漏极电压超过漏极饱和电压 V_{Dsat} 时,载流子将以极限速度或饱和速度穿过漏区附近的一部分沟道区,这个区域称为速度饱和区(VSR)。随着 V_D 的增加,速度饱和区的长度将增加,而"正常"沟道区(第 9 章的缓变沟道近似适用的区域)的长度将减小。因为沟道末端的电压保持为 V_{Dsat},而沟道长度减小,这将导致沿沟道方向的平均电场以及载流子平均速度都会增加,所以当 V_D 超过 V_{Dsat} 时,漏极电流 I_D 仍将增加。集成电路中的 MOSFET 经常偏置在载流子能达到速度饱和的工作区,因此,开发这种情形下可靠的模型是必要的。这个模型必须考虑沟道末端载流子速度达到饱和时控制沟道长度调制的物理机制。

第 8 章和第 9 章分析的基本假设是沟道中的全部可动电荷取决于 x 方向的电场,这个假设等效于将 Poisson 方程近似成一维。在缓变沟道近似中,y 方向电场的作用仅仅是使被 x 方向的电场感应出来的电荷发生移动。但在第 9 章我们已经看到,当 V_D 增加至 $V_{Dsat}=(V_G-V_T)$,并且 Q_n 减小至零时,缓变沟道模型存在矛盾。重新考虑上述基本假定,我们会注意到另一个明显的问题:漏极电压等于 $V_{Dsat}=(V_G-V_T)$ 或者更大时,一维(只和 x 有关的)Poisson 方

程是不合理的,这是因为当 $V_D > V_{Dsat}$ 并继续上升时,x 方向的电力线在夹断区(定义见第9章)中实际上已经改变了方向,一些从漏衬 pn 结空间电荷区的电离施主发出的电力线将穿过栅氧化层,并终止在栅电极上。夹断区的漏极电流在该电场的作用下,将离开 Si-SiO$_2$ 界面。为了求出这种情形下的漏极电流,必须解二维的 Poisson 方程。设 $\rho(x,y)$ 为耗尽的受主电荷浓度 qN_a 与可动电荷浓度 $-qn$ 之和,那么二维 Poisson 方程的一般形式是:

$$\frac{\partial^2 V(x,y)}{\partial x^2} + \frac{\partial^2 V(x,y)}{\partial y^2} = \frac{q[N_a(x,y) + n(x,y)]}{\epsilon_s} \tag{10.1.1}$$

式中 x 方向垂直于 Si-SiO$_2$ 界面,y 方向平行于该界面。为求解方程(10.1.1),必须先求出该区域电子浓度 n 的表达式。一般我们很难得到解析解,所以通常采用数值模拟,例如使用数值模拟软件 CADDET[1]、MINIMOS[2]或 PISCES[3]。然而,为了能从物理上理解该效应,我们宁愿采用经典方法进行分析,因此,人们发展了一种求解方程(10.1.1)的近似方法,称为"准二维模型"。

10.1.1 准二维模型

准二维模型是由 El Mansy 和 Boothroyd 在1977年[7]提出的。由于其实用性,这个模型还增加了速度饱和效应和结深的影响[8][9]。另外,该模型还可用于分析热载流子效应[10~13]。本节将推导准二维模型的基本关系式

准二维模型的速度饱和区如图10.1所示,图中的 Gauss 盒是栅极下方漏区附近的速度饱和区中空间电荷区的一部分。从 Gauss 盒子中发出的电力线沿盒子表面的积分与盒子中的电荷总量成正比,速度饱和区中的总电荷由耗尽电荷 qN_a 与可动电荷 qn 组成。Gauss 盒子一侧的边界是 xz[1] 平面(用 AB 表示),从 $\mathscr{E} = \mathscr{E}_{sat}$ 处的 Si-SiO$_2$ 界面向下延伸至漏区结深 x_j 处;另一侧也是 xz 平面(用 CD 表示),位于 y' 处,y' 的位置是可变的。设 Si-SiO$_2$ 界面处氧化层中的电场为 $\mathscr{E}_{ox}(y)$,平面 AB 为速度饱和区的起点,位于 y' 处的平面 CD 在速度饱和区中。分析中作如下假设:(1)速度饱和区中的载流子均以饱和速度运动;(2)漏结是突变 pn 结;(3)重掺杂漏区的导电性能非常好;(4)漏极电流完全限制在漏结结深 x_j 的范围内。假设(2)将分析范围限制在突变冶金结,例如掺砷的 n^+/p 结和掺硼的 p^+/n 结,但是这种方法很容易推广到缓变结,例如轻掺杂漏(LDD)结构(见10.4节)[14][15]。

图10.1中的坐标系统取 A 点处的 $y'=0$,漏区边缘的 $y'=\Delta L$。对 Gauss 盒 $ABCD$ 使用 Gauss 定理

$$-\mathscr{E}_{sat}x_j + \mathscr{E}(y')x_j + \frac{\epsilon_{ox}}{\epsilon_s}\int_0^{y'} \mathscr{E}_{ox}(y)\,\mathrm{d}y = \frac{q}{\epsilon_s}x_j y'(N_a + n) \tag{10.1.2}$$

等式(10.1.2)中使用了两个近似:(1)穿过界面 BC 的电力线与其他表面的电场积分相比可以忽略,这种假定的合理性在于,由于速度饱和区的电力线是从结平面附近发出的,因而可以近似看成是水平的;(2)y 方向的电场 $\mathscr{E}_s$ 与 x 无关(栅压 V_G 较高时近似很好,但栅压较低时近似程度较差)。将等式(10.1.2)对 y' 微分,得到

$$x_j\frac{\mathrm{d}\mathscr{E}(y')}{\mathrm{d}y'} + \frac{\epsilon_{ox}}{\epsilon_s}\mathscr{E}_{ox}(y') = \frac{q}{\epsilon_s}x_j(N_a + n) \tag{10.1.3}$$

---

1 原文误为 yz。

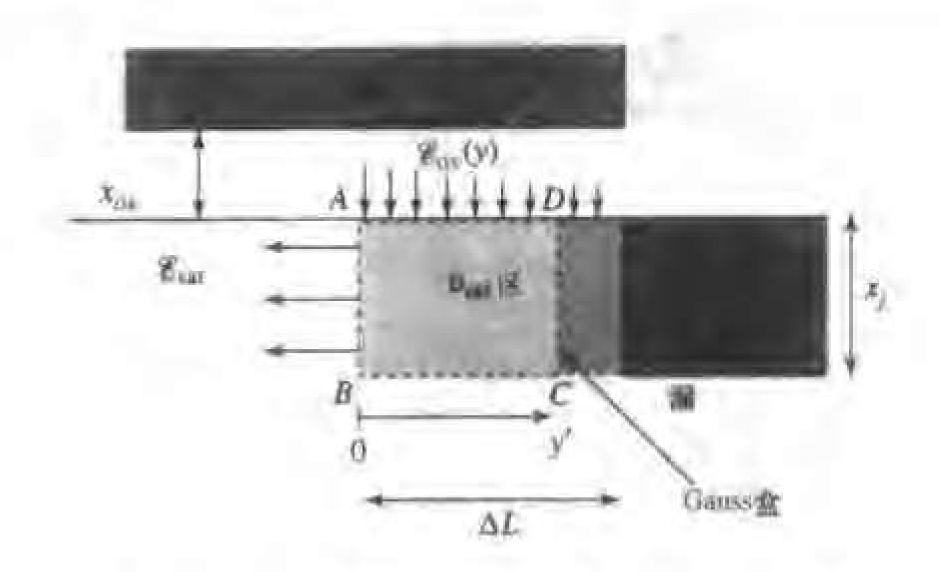


图 10.1 准二维模型的速度饱和区横截面示意图

在 Si-SiO$_2$界面处，氧化层电场 $\mathscr{E}_{ox}$ 可由下式给出

$$\mathscr{E}_{ox} = \frac{V_G - V_{FB} - 2|\phi_p| - V(y')}{x_{ox}} \tag{10.1.4}$$

将等式(10.1.4)代入(10.1.3)，得到

$$x_j \frac{\mathrm{d}\mathscr{E}(y')}{\mathrm{d}y'} + \frac{\epsilon_{ox}}{\epsilon_s}\frac{1}{x_{ox}}[V_G - V_{FB} - 2|\phi_p| - V(y')] = \frac{q}{\epsilon_s}x_j(N_a + n) \tag{10.1.5}$$

因为

$$\frac{V_G - V_{FB} - 2|\phi_p| - V_{Dsat}}{x_{ox}} = \mathscr{E}_{ox}(y' = 0) = \frac{q}{\epsilon_{ox}}x_j(N_a + n)$$

等式(10.1.5)可简化成

$$\epsilon_s x_j \frac{\mathrm{d}\mathscr{E}(y')}{\mathrm{d}y'} = C_{ox}[V(y') - V_{Dsat}] \tag{10.1.6}$$

或

$$\frac{\mathrm{d}\mathscr{E}(y')}{\mathrm{d}y'} = \frac{[V(y') - V_{Dsat}]}{\ell^2} \tag{10.1.7}$$

式中

$$\ell^2 = \frac{\epsilon_s}{\epsilon_{ox}}x_{ox}x_j \tag{10.1.8}$$

等式(10.1.8)是一个与 MOSFET 尺寸有关的参数，单位为长度的平方。

理解等式(10.1.6)的物理意义是非常有益的。沟道中 y 轴原点位置(Gauss 盒的 A 点)的电压 $V(y'=0) = V_{Dsat}$，漏极电流由以饱和速度运动的可动电荷 Q_n 组成，假设这一点上所有指向 Q_n 的电力线都是从栅极发出的(即与 $\mathscr{E}_{ox}$ 有关)。由于从 $y'=0$ 到漏移动时，硅中的电位增加到 V_D，$\mathscr{E}_{ox}$ 必然随之减小。又因为速度饱和区中的 I_D 必须是连续的，Q_n 也应该是常数。Gauss 盒中要满足这些条件，只能是电场从漏区向速度饱和区增加，并且增加的数值与 $\mathscr{E}_{ox}$ 减小的数值相等，那么由 Possion 方程，氧化层电场减小而释放的电荷必须被 y' 方向增加的电场维持。等式(10.1.6)的右端是沟道电压增量为 $V(y') - V_{Dsat}$ 而导致的氧化层电场释放出的电荷总量，左端对应于为维持这些电荷量而引起的沟道电场梯度的增量。由氧化层电场和漏电场分别控制的电荷如图 10.2 所示。

利用边界条件 $\mathscr{E}(0) = \mathscr{E}_{sat}$ 和 $V(0) = V_{Dsat}$，求解等式(10.1.7)，得到

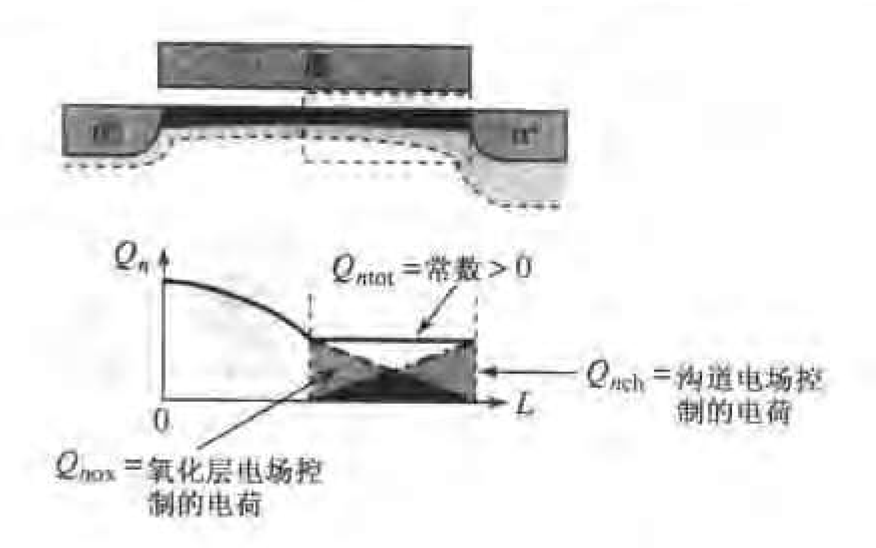


图 10.2　准二维近似中的由氧化层电场和漏电场分别控制的电荷

$$\mathscr{E}(y') = \mathscr{E}_{sat}\cosh(y'/\ell) \tag{10.1.9}$$

和

$$V(y') = V_{Dsat} + \ell\,\mathscr{E}_{sat}\sinh(y'/\ell) \tag{10.1.10}$$

在漏端，沟道电场强度达到最大值

$$\mathscr{E}_m = \mathscr{E}(y' = \Delta L) = \mathscr{E}_{sat}\cosh(\Delta L/\ell) \tag{10.1.11}$$

和

$$V_D = V_{Dsat} + \ell\,\mathscr{E}_{sat}\sinh(\Delta L/\ell) \tag{10.1.12}$$

联立等式(10.1.11)和(10.1.12)，可求出速度饱和区的长度

$$\Delta L = \ell\ln\left\{\frac{[(V_D - V_{Dsat})/\ell] + \mathscr{E}_m}{\mathscr{E}_{sat}}\right\} \tag{10.1.13}$$

以及沟道中的最大电场强度

$$\mathscr{E}_m = \left[\frac{(V_D - V_{Dsat})^2}{\ell^2} + \mathscr{E}_{sat}^2\right]^{1/2} \tag{10.1.14}$$

从等式(10.1.9)到(10.1.14)可以得到一些重要的物理概念。等式(10.1.9)表明沟道电场沿 y' 方向几乎是指数增长的，而为了维持氧化层电场释放的电荷，这种陡峭的增长是必要的。用等式(10.1.9)计算出的沟道电场 $\mathscr{E}(y')$ 如图 10.3(a)所示，与图 10.3(b)所示的二维数值模拟结果非常相似。参数 ℓ 被调整到使图 10.3(a)和图 10.3(b)中的最大沟道电场强度近似相等。

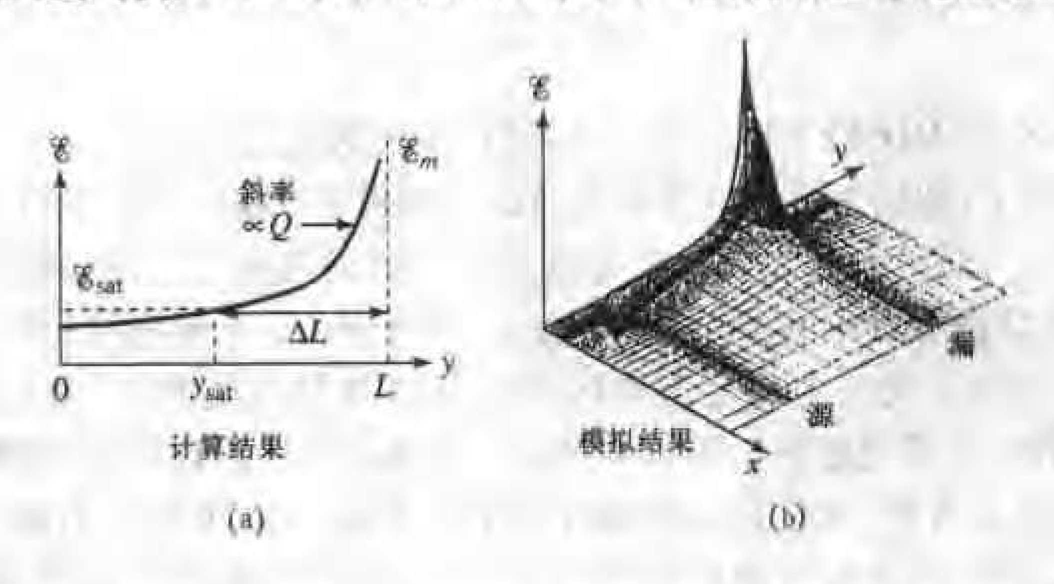


图 10.3　靠近 $Si\text{-}SiO_2$ 界面处硅一侧的电场分布：(a)准二维近似计算结果；(b)二维数值模拟结果。注意，表面速度饱和区 ΔL 中靠近漏端的电场增长非常陡峭

为了研究短沟 MOSFET 的最大沟道电场,我们在等式(10.1.10)和(10.1.14)中引入一些实际的数据。对于典型的 1μm 工艺 n 沟道 MOSFET,取 $x_{ox}=25\text{nm}$,$x_j=0.2\text{nm}$,由等式(10.1.8)求出 $1/\ell=8\times10^4\text{V/cm}$。当$(V_G-V_T)$很小时,$V_{Dsat}$也很小,可忽略沟道电场的影响。取$\mathscr{E}_{sat}=5\times10^4\text{V/cm}$,$V_D=5\text{V}$,由等式(10.1.14)求出$\mathscr{E}_m=4.0\times10^5\text{V/cm}$,这是能引起多种热载流子效应的强电场。因为 $\mathscr{E}_m$正比于 $1/(x_{ox}^{1/2}x_j^{1/2})$,随着 MOSFET 尺寸的减小,强场效应将更为严重。将电源电压减小至 3.3V 有助于缓和强场效应。但是为了提高器件的可靠性,我们仍然需要能抗热载流子效应的器件结构,这将在 10.4 节讨论。

最大沟道电场 $\mathscr{E}_m$ 显然与(V_D-V_{Dsat})有关,但 MOSFET 的沟道长度只是间接通过 V_{Dsat}来影响$\mathscr{E}_m$。当漏极电压 V_D比 V_{Dsat}高出 1V 或 2V 时,$(V_D-V_{Dsat})/\ell\gg\mathscr{E}_{sat}$,那么最大电场强度近似正比于$(V_D-V_{Dsat})$,并可以表示为:

$$\mathscr{E}_m\approx\frac{(V_D-V_{Dsat})}{\ell},\quad 当(V_D-V_{Dsat})>1\text{V}时, \tag{10.1.15}$$

二维数值分析证实了准二维分析的基本方程(10.1.9)和(10.1.13),这说明虽然存在近似,但准二维分析保留了强场效应的物理本质。我们注意到衬底掺杂浓度 N_a没有出现在上述任何表达式中,二维数值模拟和实验结果都表明 N_a对 ℓ 几乎没有影响。N_a的大部分影响是通过 V_{Dsat}反映出来的,而这一影响在短沟器件中较小。二维模拟表明 ℓ 对 x_{ox}的依赖关系比等式(10.1.8)描述的要弱

$$\ell=0.22x_j^{1/2}x_{ox}^{1/3} \tag{10.1.16}$$

式中 ℓ、x_j和 x_{ox}的单位都是 cm。等式(10.1.16)适用于较大范围的 x_j、x_{ox}和 N_a,因此得到广泛使用。数值模拟与准二维模型相比,ℓ 与 x_{ox}的关系稍微减弱了一些,由 1/2 次幂变为 1/3 次幂。

例题 最大电场

绘出最大横向电场 $\mathscr{E}_m$与 V_D的关系曲线。

解:V_D大于或小于 V_{Dsat}这两种情况下,$\mathscr{E}_m$表达式有不同的形式。如果 $V_D<V_{Dsat}$,器件工作在线性区,由等式(9.2.8)得

$$I_D=C_{ox}W[V_G-V_T-V(y)]\frac{\mu_{eff}\mathscr{E}(y)}{1+[\mathscr{E}(y)/\mathscr{E}_{sat}]}$$

将电场表示成负的微分电压,并积分到任意位置 y,得到

$$I_Dy=\mu WC_{ox}\left(V_G-V_T-\frac{V(y)}{2}\right)V(y)\frac{1}{1+(V_D/\mathscr{E}_{sat}y)}$$

经过适当的整理,解出 $V(y)$

$$V(y)=(V_G-V_T)-\sqrt{(V_G-V_T)^2-\frac{2[1+(V_D/\mathscr{E}_{sat}y)]I_Dy}{\mu WC_{ox}}}$$

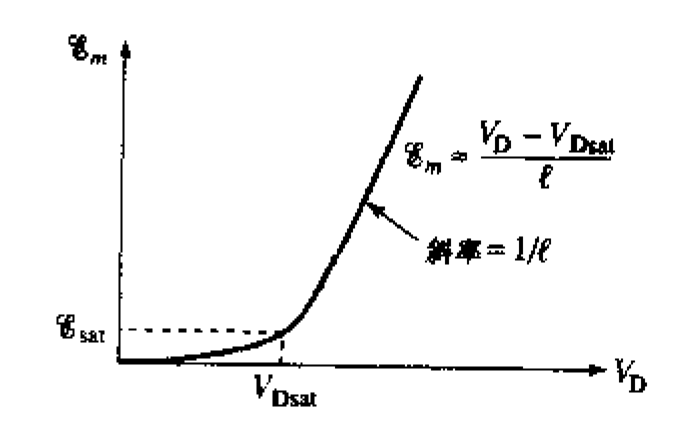
电场 $\mathscr{E}(y)$为

$$\mathscr{E}(y)=-\frac{dV(y)}{dy}=\frac{I_D}{\mu WC_{ox}\sqrt{(V_G-V_T)^2-\frac{2[1+(V_D/\mathscr{E}_{sat}y)]I_Dy}{\mu WC_{ox}}}}$$

$\mathscr{E}_m$为 $y=L$ 处的电场

$$\mathscr{E}_m(V_D)=\frac{\dfrac{\mu WC_{ox}}{L}\left(V_G-V_T-\dfrac{V_D}{2}\right)V_D\dfrac{1}{1+(V_D/\mathscr{E}_{sat}L)}}{\sqrt{[\mu WC_{ox}(V_G-V_T)]^2-2(\mu WC_{ox})^2\left(V_G-V_T-\dfrac{V_D}{2}\right)V_D}}$$

如果 $V_D>V_{Dsat}$,则$\mathscr{E}_m$由等式(10.1.14)给出,由等式(10.1.15)可得到$\mathscr{E}_m$的近似解,相应的$\mathscr{E}_m$ ~ V_D曲线如下图所示



利用等式(10.1.13)中 ΔL 的表达式,可以推导出饱和区漏极电流和等效输出电阻 R_{out}的表达式,$V_D>V_{Dsat}$时,漏极电流为

$$I_D=I_{Dsat}\left[\frac{V_G-V_T+\mathscr{E}_{sat}L}{V_G-V_T+\mathscr{E}_{sat}(L-\Delta L)}\right] \tag{10.1.17}$$

等效输出电阻 R_{out}(I_D ~ V_D曲线斜率的倒数)可以利用等式(10.1.17)对 V_D微分得到,经过一系列数学处理,得到 R_{out}的近似解为

$$R_{out}=\frac{1}{I_{Dsat}}\left(\frac{V_D-V_{Dsat}}{\ell}\right)\left(L+\frac{V_G-V_T}{\mathscr{E}_{sat}}\right) \tag{10.1.18}$$

由等式(10.1.18)给出的有限大小的输出电阻会使电路性能退化。在理想情况下我们希望由源极提供的漏极电流只由栅极电压控制,而与负载无关。在 MOSFET 不存在严重的泄漏电流或者穿通的条件下,等式(10.1.18)指出 R_{out}与 $1/I_{Dsat}$成正比,这种函数关系与观测到的结果非常接近。图 10.4 给出了测量得到的饱和区输出电阻与漏极电流的关系,其中沟道长度 L 和氧化层厚度 x_{ox}与文献[16]中的参数一致。这些数据验证了输出电阻与 L 和 $1/x_{ox}^n$的关系(其中 $1/3<n<1/2$),但参数 ℓ 与结深 x_j的关系很难通过实验得到验证,因为 R_{out}不仅与 x_j的大小有关,还与结的结构有关,我们很难制造 x_j改变但结的结构不变的器件。但是,二维分析有助于证明等式(10.1.16)和(10.1.18)中与 ℓ 的关系。

本节给出的准二维模型提供了一种有用的分析方法,可以确定 MOSFET 速度饱和区中最大电场强度的近似值。下面几节首先将考虑一些对器件性能有害的强场效应,然后介绍一些利用强场效应的特殊器件。

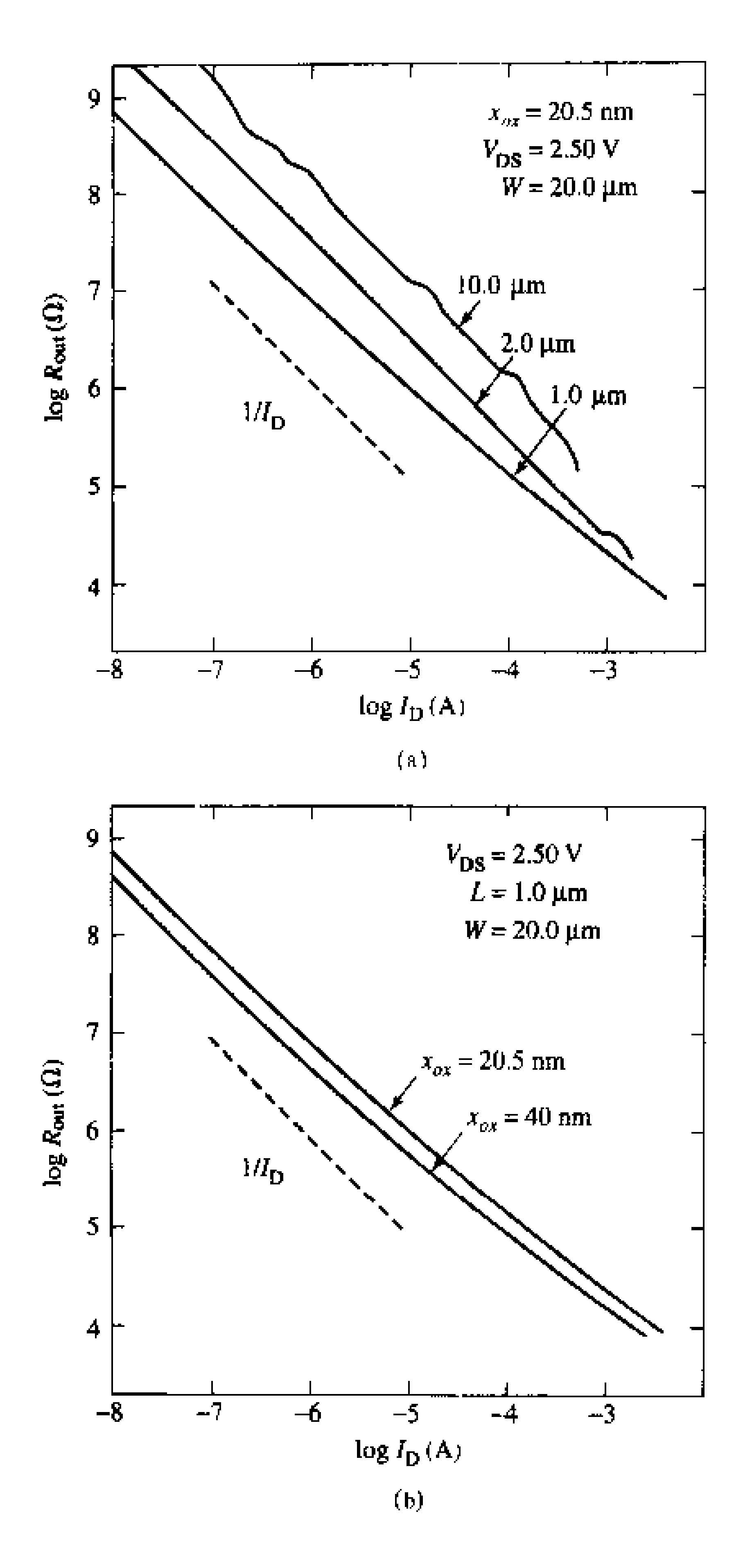


图10.4 实验测得的饱和区输出电阻与漏极电流的关系：(a)沟道长度 L 改变；(b)氧化层厚度 x_{ox} 改变 图中输出电阻与漏极电流近似的倒数关系非常明显[16](© 1987 IEEE)

10.2 衬底电流

10.2.1 热载流子效应

当MOSFET的尺寸减小到远小于1μm(通常称为深亚微米)时，遇到的最严重的问题之一

是热载流子效应。例如,1V 的电势差加在0.5μm 长度上,产生的平均电场为 2×10^4 V/cm,这已经足以使电子和空穴的速度达到饱和。在这样的电场下,载流子被称为"热"载流子,它们的高能量导致的物理效应能使 MOSFET 性能退化。图 10.5 举例说明了热载流子效应产生的寄生电流(通常是不需要的)的主要成分。衬底电流若不加以限制,会引起局部的电势波动,或者向衬底注入电子,诱发"snap-back"器件击穿(电压随电流的升高而降低),还会引起 9.3 节讨论的 CMOS 闩锁效应。电子注入氧化层引起的栅极电流能够使氧化层性能退化,但也提供了一种制造可编程不挥发器件的方法,这将在 10.5 节讨论。总之,热载流子效应不仅有害,还可被利用,所以了解热载流子效应是非常重要的。

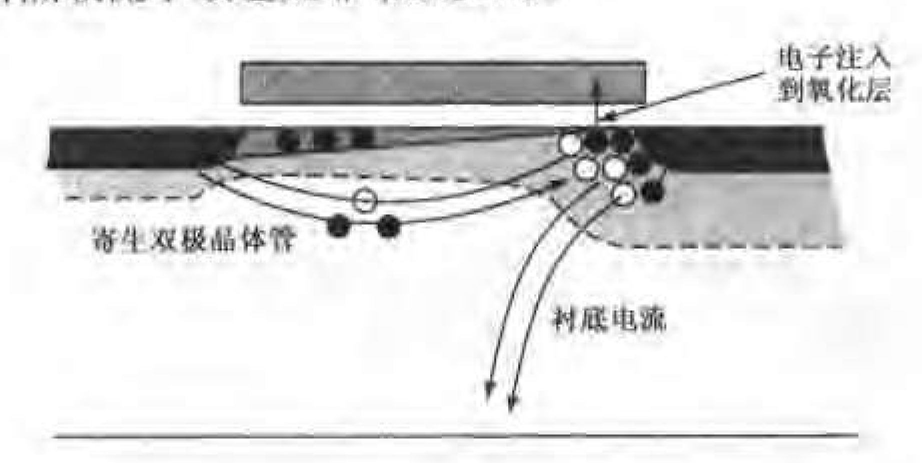


图 10.5 MOSFET 中沟道热载流子引起的有害的效应:载流子被注入到栅氧化层中;电流流向衬底;寄生双极晶体管进入正向放大区

当电场强度超过 $3\sim5\times10^4$ V/cm 时,电子在反型层中的速度将达到饱和。热载流子效应在电场低于上述值的区域一般不重要,所以只需要考虑速度饱和区。热载流子效应最重要的监控参数是最大沟道电场 $\mathscr{E}_m$。10.1 节已经给出了$\mathscr{E}_m$的表达式(10.1.14(和(10.1.15),$\mathscr{E}_m$的大小主要由(V_D-V_{Dsat})和 ℓ 决定。由上述理论可知,当漏极电压 V_D升高,沟道长度 L 减小(导致 V_{Dsat}减小),栅氧化层减薄(导致 ℓ 减小),结深减小(导致 ℓ 减小)时,热载流子效应将更严重。热载流子效应产生的主要电流成分是衬底电流和栅极电流,下面分别讨论。

10.2.2 衬底电流模型

如果电子能在 MOSFET 的沟道中获得大于 1.5eV 的动能,它们与晶格碰撞时就能导致碰撞电离,产生电子-空穴对。碰撞电离产生的电子或者被吸引到势能较高的漏极(加入到漏极电流中),或者注入到栅氧化层中(如果它们具有足够的能量)。而产生的空穴却与之相反,被吸引到衬底(势能较低的位置),产生寄生衬底电流 I_{sub}。尽管这种载流子产生模式能导致雪崩击穿,但是由于产生的空穴能很快地移动到低场区,雪崩倍增并不经常发生。为了正确估算 n 沟道 MOSFET 衬底电流值的大小,需要计算速度饱和区中碰撞电离产生的空穴数。

一个载流子在单位长度内发生的碰撞电离的次数可以用电离率来计算。由等式(4.4.13)知,碰撞电离率是随电场强度变化的,与电场强度的关系为 $A_i\exp(-B_i/\mathscr{E})$,其中电子的电离参数 B_i大约为空穴的两倍,这使得空穴导致的碰撞电离没有电子的碰撞电离重要。考虑 n 沟道 MOSFET,如果假设所有的空穴都流向衬底,那么由于漏极电流碰撞电离产生的衬底电流表达式为[17]

$$I_{sub}=\int_0^{\ell} I_D A_i\exp[-B_i/\mathscr{E}(y)]\,dy \tag{10.2.1}$$

将等式(10.2.1)中的 dy 用(dy/d$\mathscr{E}$)d$\mathscr{E}$或 $-\mathscr{E}^2$(dy/d$\mathscr{E}$)d(1/$\mathscr{E}$)替换,得到

$$I_{sub} = \int_{\mathscr{E}_s}^{\mathscr{E}_m} -I_D A_i \exp[-B_i/\mathscr{E}(y)]\ \mathscr{E}^2(y)\ \frac{dy}{d\mathscr{E}} d\left(\frac{1}{\mathscr{E}}\right) \qquad (10.2.2)$$

式中$\mathscr{E}_s$为沟道源端电场强度。利用等式(10.1.9),并将 cosh(y/ℓ)近似成 $e^{y/\ell}/2$,得到

$$\frac{d\mathscr{E}}{dy} \approx \mathscr{E}_{sat}\frac{e^{y/\ell}}{2\ell} = \frac{\mathscr{E}}{\ell} \qquad (10.2.3)$$

因此

$$\mathscr{E}^2(y)\ \frac{dy}{d\mathscr{E}} \approx \ell\mathscr{E} \qquad (10.2.4)$$

由于等式(10.2.2)中的指数项显然在$\mathscr{E}=\mathscr{E}_m$处达最大值,我们取电场值$\mathscr{E}=\mathscr{E}_m$,并设在积分区间为常数,这样$\mathscr{E}_m$就可以移出到积分号外。将等式(10.2.4)代入等式(10.2.2),经过这样的简化后[2],得到 I_{sub}为

$$I_{sub} = A_i \ell \mathscr{E}_m I_D \int_{\mathscr{E}_s}^{\mathscr{E}_m} \exp[-B_i/\mathscr{E}(y)]d\left(\frac{1}{\mathscr{E}}\right)$$

$$= \frac{A_i}{B_i}\ell \mathscr{E}_m I_D \exp(-B_i/\mathscr{E})\Big|_{\mathscr{E}_s}^{\mathscr{E}_m} \approx \frac{A_i}{B_i}\ell \mathscr{E}_m I_D \exp(-B_i/\mathscr{E}_m) \qquad (10.2.5)$$

可近似为

$$I_{sub} \approx \frac{A_i}{B_i}(V_D - V_{Dsat}) I_D \exp\left(-\frac{\ell B_i}{V_D - V_{Dsat}}\right) \qquad (10.2.6)$$

等式(10.2.5)的这种形式与前面的陈述是一致的,即衬底电流主要决定于($V_D - V_{Dsat}$)以及参数 ℓ。实验表明 B_i近似等于 1.7×10^6 V/cm,A_i/B_i近似为 1.2V^{-1}

在继续讨论衬底电流之前,有必要对它的大小进行说明。电子从外加电场获得的总能量,等于相邻两次碰撞间作用在电子上的总库仑力沿距离的积分。用平均自由程乘以电子所处的电场,这种平均的方法计算热电子能量的正确性还没有得到证实。但是由于某些推导,我们必须考虑这一问题,例如,碰撞电离率的表达式(等式(4.4.13))就在很大程度上取决于电场的大小。当电场随位置变化很快时,电子能量将比通过局部电场和平均自由程估算出来的值小如果在电离率表达式中采用碰撞点的电场,又会高估碰撞电离率。完整的分析需要用复杂的数值计算,而不是解析方法。然而,数学方法往往会限制对物理的洞察,所以近似分析仍然是有意义的。

在 $I_D \sim V_D$特性曲线图中叠加 I_{sub}/I_D曲线,可以得到关于 I_{sub}的一些有用的信息。如图 10.6 所示,I_{sub}/I_D的曲线族彼此平行,每条曲线不仅描述了 I_{sub}/I_D,还描述了$\mathscr{E}_m$(见等式(10.2.5)),因此也描述了($V_D - V_{Dsat}$)(见等式(10.2.6))。图 10.6 证实了 I_{sub}/I_D曲线族平行于 V_{Dsat}曲线[18]

图 10.7 为(I_{sub}/I_D)与 1/($V_D - V_{Dsat}$)关系的测量曲线,其中利用了等式(10.2.6)的形式,MOSFET 的沟道长度分别为 0.95μm、1.45μm 和 2.7μm,栅极电压分别为 0.9V 和 1.5V。这些不同的栅极电压和不同的沟道长度下的所有数据点都落在同一条直线上,这个线性关系证明了推导等式(10.2.6)所用的假设,也证明了衬底电流与栅极电压以及沟道长度的关系是通过 V_{Dsat}项来体现的。从等式(10.2.5)推导等式(10.2.6)时,假设$\mathscr{E}_{sat}$远小于$\mathscr{E}_m$并忽略了$\mathscr{E}_{sat}$。

---

2 译者注:推导等式(10.2.5)时,只是将等式(10.2.4)描述的项移出到积分号外,其余包含 ℓ 的两项并没有移出到积分号外

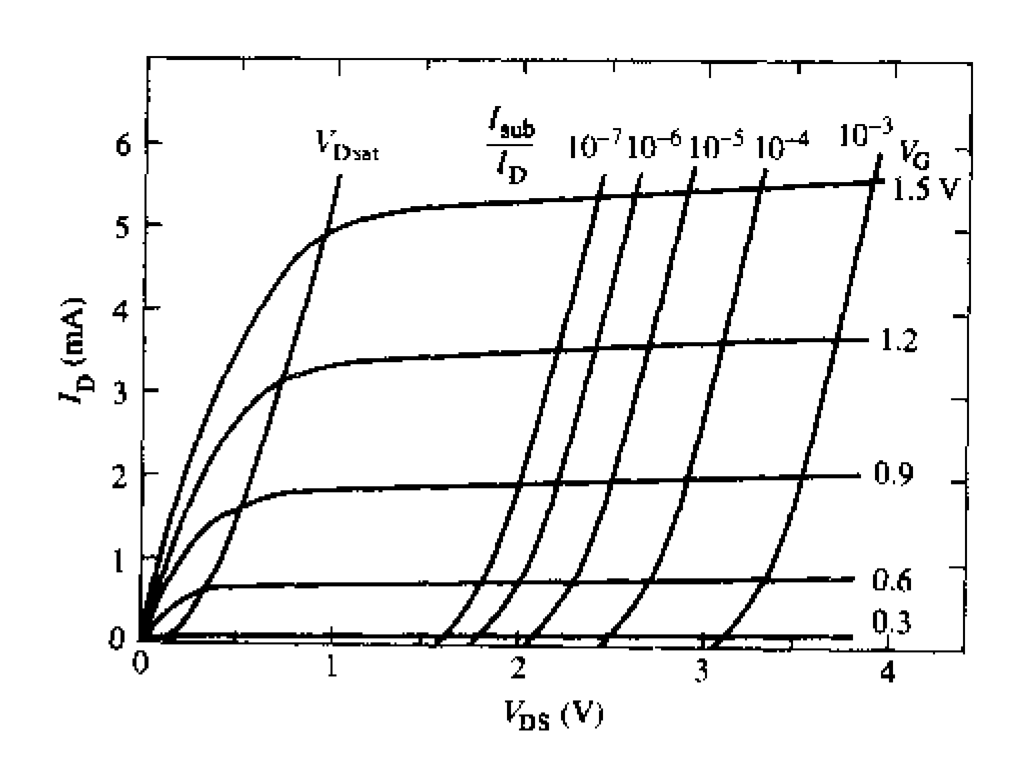


图 10.6 $x_{ox}=15.2$nm, $L=1.45\mu m$ 的 n 沟道 MOSFET 的 $I_D \sim V_D$ 曲线以及 I_{sub}/I_D 曲线

实验数据表明图 10.7 中的线性关系在 I_{sub}/I_D 小到 10^{-6} 时仍然有效,当低于这个数值时$\mathscr{E}_{sat}$不再远小于$\mathscr{E}_m$。我们注意到图 10.7 中曲线的斜率是 $-\ell B_i$,由此可以获得 B_i 的值。由等式(10.2.6)知,通过曲线与纵轴的交点可以得到 A_i/B_i 的值。等式(10.2.6)对于 $B_i\approx3.7\times10^{-6}$V/cm, $A_i/B_i=2.2$ 的 p 沟道 MOSFET 也是有效的。

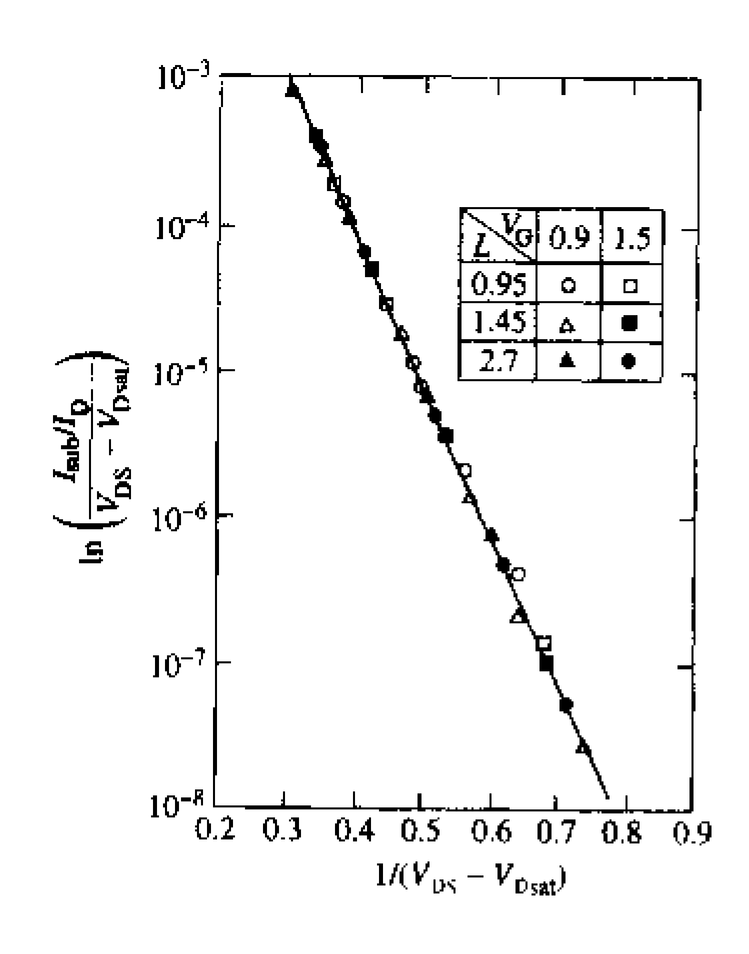


图 10.7 测量得到的(I_{sub}/I_D)与 $1/(V_D-V_{Dsat})$ 的关系曲线,其中 MOSFET的沟道长度分别为 0.95μm、1.45μm 和 2.7μm,栅极电压分别为 0.9V 和 1.5V。纵轴的形式利用了等式(10.2.6)(© 1984 IEEE)

尽管图 10.7 中的曲线能有效地验证等式(10.2.6),但它并没有直接给出 I_{sub} 与端电压的关系。图 10.8 是以 V_G 为变量的 I_{sub} 曲线,其中 V_D 取了五个不同的值。该曲线更为直接地反映了 I_{sub} 与外加电压的关系[9],并表明 V_G 在 V_T 和 $V_D/2$ 之间时,I_{sub} 达到极大值。该现象可以这样

从物理上解释：首先，在较低的栅压下，对于一定的碰撞电离率，速度饱和区必须存在一定数量的反型电子，这些电子输运漏极电流 I_D。等式(10.2.6)表明，I_{sub}和 I_D之间存在比例关系，因此，增大栅压往往导致漏极电流提高而使 I_{sub}增加。但是，增大栅压的同时还增大了 V_{Dsat}，由等式(10.1.14)知，这将减小最大电场 $\mathscr{E}_m$。等式(10.2.5)表明，$\mathscr{E}_m$的减小将导致衬底电流 I_{sub}的下降。这两种机构共同作用的结果是，$I_{sub} \sim V_G$曲线上存在如图 10.8 所示的极值点。

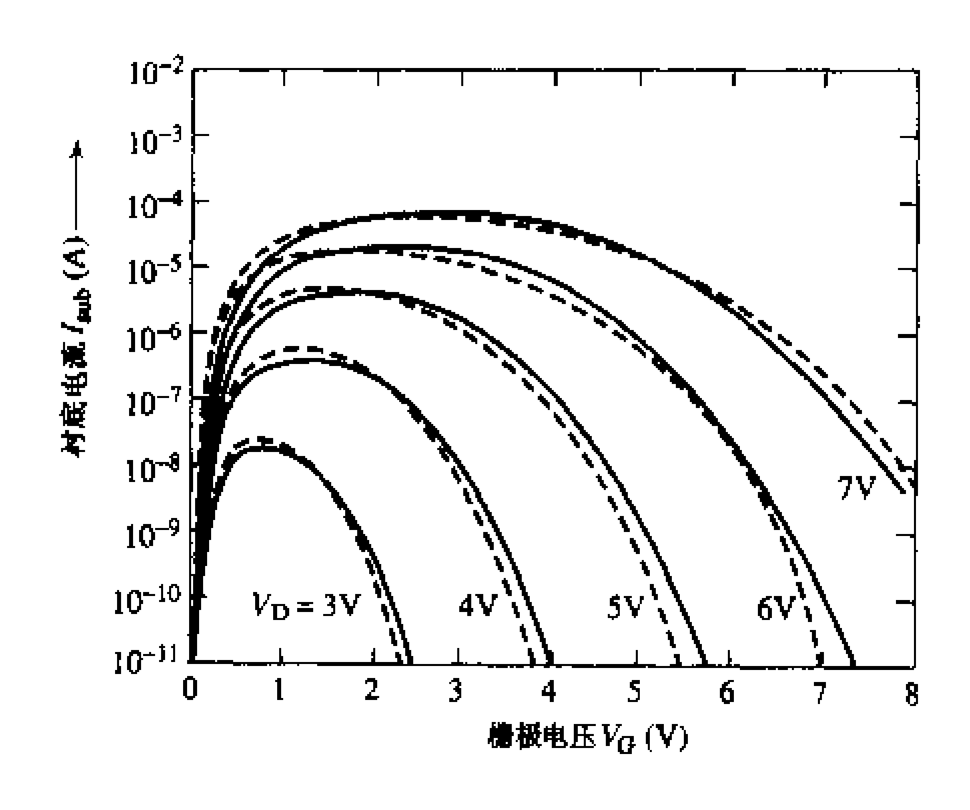


图 10.8　在不同的漏极电压 V_D下，衬底电流 I_{sub}与栅极电压 V_G的关系曲线。实线为测量结果，虚线为计算结果

以 V_G为参量，衬底电流 I_{sub}与漏极电压 V_D的关系如图 10.9 所示。首先在($V_D - V_{Dsat}$) ≈ 1V 时(类似于从等式(10.1.14)推导等式(10.1.15)[3]的条件)观测到较大的衬底电流 I_{sub}，漏极电压增加时，衬底电流线性增加。图 10.9 中曲线的典型斜率近似为漏极电压每增加 0.5V，衬底电流大约增加 10 倍。更高的栅压 V_G将导致更高的 V_{Dsat}，从图 10.9 中可看出，不同 V_G下的 I_{sub}曲线沿 V_D轴移动 V_{Dsat}的改变量。这种现象是合理的，因为 I_{sub}不直接与 V_D有关，而是依赖于 V_D与 V_{Dsat}的差。

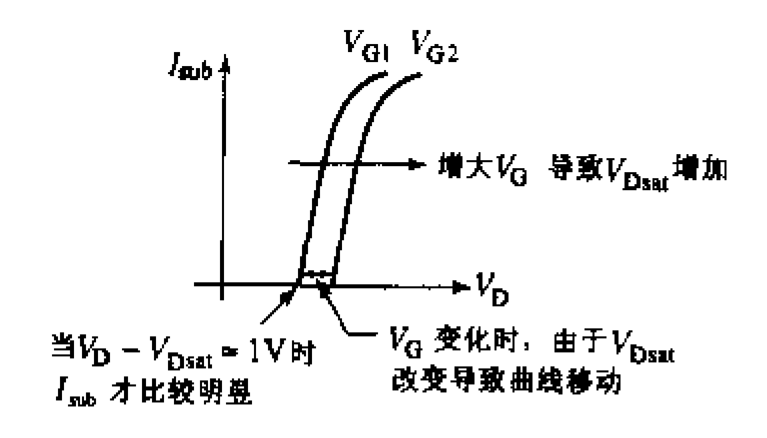


图 10.9　I_{sub}与 V_D的关系曲线。在两个不同的栅压 $V_{G2} > V_{G1}$下，因为 V_{Dsat}增加而使曲线向右移动

[3] 译者注：原文误为(10.1.13)。

10.2.3　衬底电流对漏极电流的影响

MOSFET 的衬底电流 I_{sub} 流经衬底时，会在衬底电阻上产生电压降，如图 10.10 所示。这个电压降将提高沟道下方中性区的电位，产生反衬偏效应，使 MOSFET 的阈值电压减小、漏极电流增加。因为衬底电流 I_{sub} 与漏极电压 V_D 有关，所以这种现象也是一种漏极电压调制饱和区的漏极电流的效应，因而也减小了输出电阻。

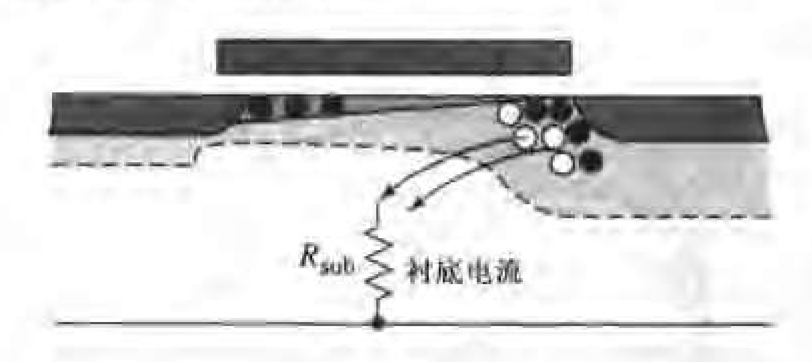


图 10.10　衬底电流经衬底流向衬底电极，使外加的衬偏电压与沟道下方中性区的电位不同，因此，有效的衬偏电压改变，导致晶体管的阈值电压也发生了变化

前面已经讨论了很多效应，因此有必要归纳影响饱和漏极电流（和 R_{out}）的重要因素：(1)沟道长度调制(CLM)，(2)漏致势垒降低(DIBL)，(3)衬底电流导致的体效应(SCBE)[20]。前面在讨论沟道长度调制时认为 R_{out} 与 $(V_D - V_{Dsat})$ 成正比，这就是等式(10.1.18)。然而，实验得出的结论是，R_{out} 在 V_D 略高于 V_{Dsat} 时达到极大值，然后随 V_D 的增加迅速下降到非常低的值，如图 10.11 所示。R_{out} 的快速下降现象出现在热载流子导致的衬底电流起主要作用时。漏致势垒降低（见 9.2 节）除了使 R_{out} 减小外，还会导致强反型晶体管的阈值电压随 V_D 的增加而减小，但是通过合理的 MOSFET 设计，可以使阈值电压的下降最小化。CLM、DIBL 和 SCBE 效应都能使饱和漏极电流增加，输出电阻下降，它们的相对作用强烈依赖于偏置条件，如图 10.11 所示。准确描述衬底电流和输出电阻的模型非常复杂[20]，本文不涉及这种定量讨论。

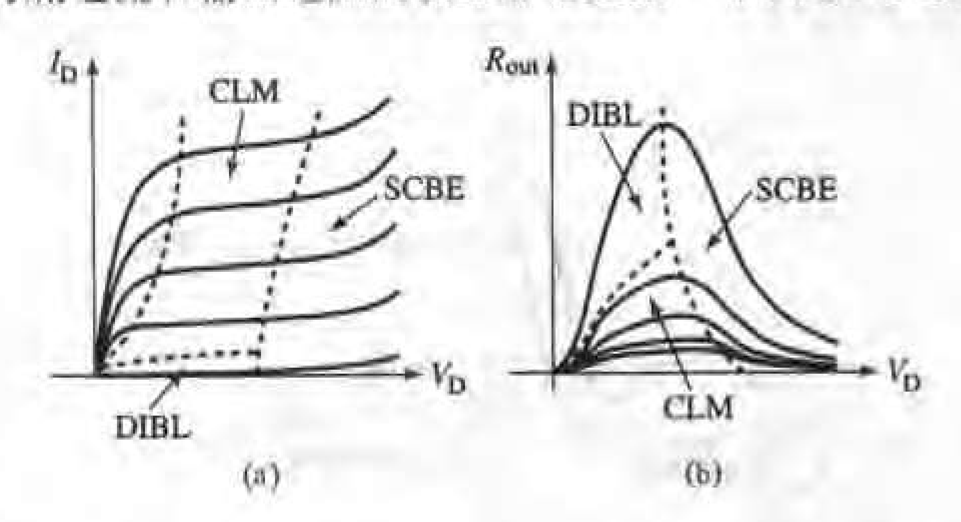


图 10.11　由沟道长度调制(CLM)、漏致势垒降低(DIBL)以及衬底电流导致的体效应(SCBE)引起的(a)漏极电流增加和(b)输出电阻减小。图中还描述了各效应起主要作用的偏置区域

当 MOSFET 衬底的电流增加到使沟道下方的电位超过源-衬结的导通电压时，这个结将起到双极晶体管发射结的作用，此时沟道下方的中性区相当于基区，漏区相当于集电区。这个寄

生 npn 双极晶体管如图 10.12 所示。由于源-衬结正偏，寄生双极晶体管偏置在放大区，导致电子从源区注入到衬底区，这将进一步提高衬底电位，减小 MOSFET 的阈值电压，从而使沟道电流增加，产生更多的碰撞电离（更高的 I_{sub}），最终将导致晶体管击穿，如图 10.13 所示。由于这些效应，MOSFET 的击穿电压强烈依赖于衬底电流的大小和衬底有效电阻。图 10.13 给出了经常能观测到的 C-型击穿曲线（虚线），这意味着最低的击穿电压出现在中等栅压区。临界栅压与 I_{sub}的大小有关，而 V_G在 V_T 和 $V_D/2$ 之间时 I_{sub}达到最大值，如图 10.8 所示。

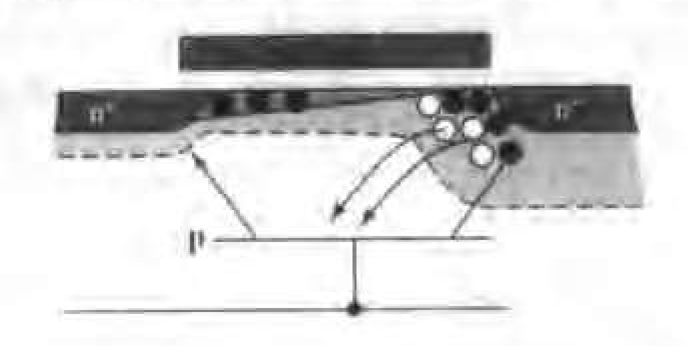


图 10.12　当 MOS 晶体管的源-衬结正偏时，这个结将起到寄生双极晶体管的发射结的作用，漏区相当于寄生双极晶体管的集电区

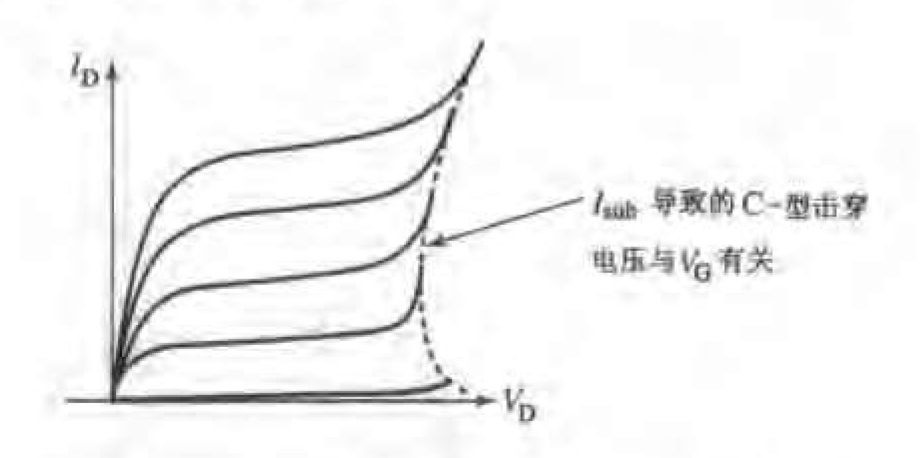


图 10.13　MOSFET 沟道下方的寄生双极晶体管导致的 C-型击穿现象

10.3　栅极电流

前面已经讨论了硅中不同区域的电流，并假定栅氧化层是理想的绝缘体。本节将讨论栅极电流，它是由沟道中的载流子注入到氧化层，并穿过氧化层流向栅电极引起的，其中最重要的导电机制是高能（热）电子向氧化层的注入。热载流子注入意味着载流子已不再与注入点的晶格保持平衡，也就是说，热载流子具有的动能远高于晶格温度对应的能量。载流子在沟道中被强电场加速并获得能量，当它们的能量足以越过图 8.2 所示的 Si-SiO_2 界面势垒时，就会发生热载流子注入。热载流子注入与隧道注入不同，隧道注入是指与硅晶格保持热平衡的载流子（即"冷"载流子）穿过很薄的氧化层。

10.3.1　幸运电子模型

尽管衬底电流与栅极电流都是由热载流子效应引起的，但与之相关的偏置电压却完全不同。图 10.14 给出了观测到的 n 沟道 MOSFET 的栅极电流和衬底电流[21]。从图中可看出，栅

极电流远小于衬底电流,并且当栅极电压和漏极电压接近相等时,栅极电流达到最大值。沟道中的热电子能够到达栅极,必须满足两个条件:(1)热电子必须在沟道电场中获得足够的动能,以克服 $Si-SiO_2$ 界面势垒[11];(2)热电子必须改变运动方向,指向 $Si-SiO_2$ 界面[22]。图 10.15 描述了这两个必要条件:当电子在沟道中从 A 点向 B 点运动时,获得动能,并在 B 点运动方向改变为指向氧化层(C 点)。热电子在改变运动方向时,碰撞损失的能量必须很小,以保证它们有足够的能量克服 $Si-SiO_2$ 界面的势垒。一旦从硅中发射的电子越过势垒进入氧化层,只要栅极电势高于硅表面的沟道电势,就会被氧化层中的电场扫向栅电极。从统计学的角度来看,这两个过程都是相互独立的,因此总的概率等于每个独立事件概率的乘积。

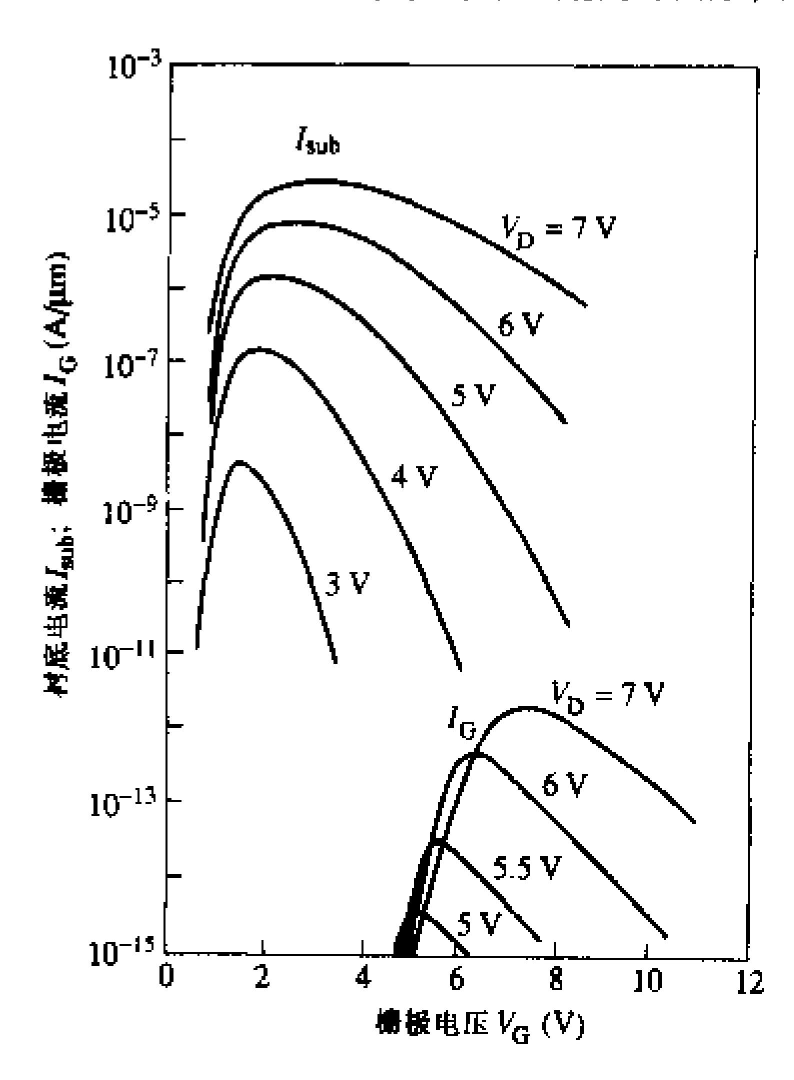


图 10.14 n 沟道 MOSFET 在不同的 V_D 和 V_G 下测得的衬底电流与栅极电流(A/μm) 该图表明,与 I_G 相比,I_{sub} 更大,并且在更低的 V_G 下达到最大值[2] (© 1982 IEEE)

电子越过 $Si-SiO_2$ 界面势垒需要的动能为 $q\phi_B = 3.1$ eV(见图 8.2),因此形成栅极电流的热电子的动能必须高于 $q\phi_B$。如果假设沟道加速电场($\mathscr{E}_y$)是恒定的,那么热电子必须运动一段距离 $d = (\phi_B/\mathscr{E}_y)$ 以获得足够的动能。沟道电子运动这么长或者更长的距离而不发生碰撞的概率可以表示成 $\exp(-d/\lambda)$,其中 λ 是热电子的平均自由程。因此,电子获得高于 $Si-SiO_2$ 界面势垒的动能的概率为 $\exp[-\phi_B\phi/(\mathscr{E}_y\lambda)]$。这种计算栅极电流的方法通常被称为幸运电子模型[11]。如图 10.15 所示,注入和被栅电极收集的概率 P 是氧化层电场 $\mathscr{E}_{ox}$ 的函数。我们可以利用载流子数目(用漏极电流表示)、载流子获得足够能量的概率(表示成 $\exp[-\phi_B/\mathscr{E}_y\lambda]$)、发射和收集概率 P 这三项的乘积来推导栅极电流 I_G 的表达式。将这个乘积在整个沟道中积分,得到

$$I_G = I_D\int_0^l \exp\left(-\frac{\phi_B}{\mathscr{E}_y\lambda}\right)P(\mathscr{E}_{ox})\,dy \tag{10.3.1}$$

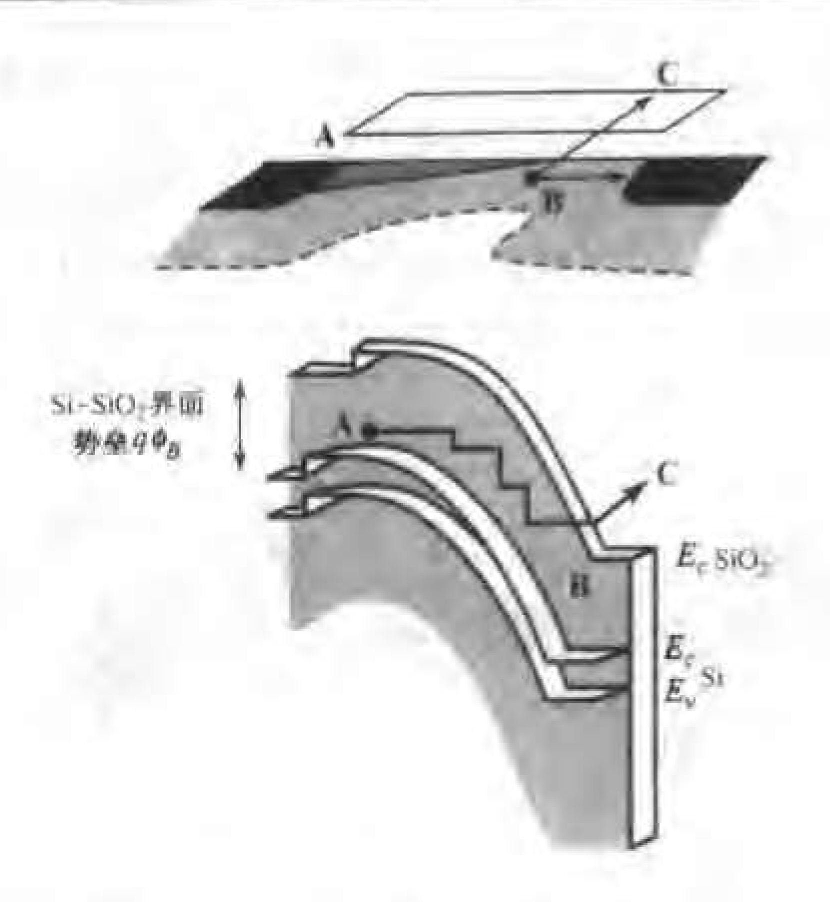


图10.15 形成栅极电流的过程:(1)载流子从A点运动到B点的过程中,必须获得足够的克服Si-SiO$_2$界面势垒的能量;(2)载流子(B点)的运动方向必须改变为指向Si-SiO$_2$界面,然后从B点进入二氧化硅的导带,并最终到达栅电极

上式可近似成

$$I_G \approx CI_D\exp\left(-\frac{\phi_B}{\lambda\mathscr{E}_m}\right) \tag{10.3.2}$$

当$V_G > V_D$时,C近似等于2×10^{-3}。

下面的定性讨论会经常地比较栅极电压和漏极电压的大小。为了描述方便,我们假定阈值电压$V_T=0$(也就是说,可以认为有效栅压V_G-V_T正好等于V_G),并将源极电压作为参考点。

如图10.16所示,对于一个确定的V_G值,因为最大电场强度$\mathscr{E}_m$随漏极电压增大,栅极电流会随漏极电压的增加而增大,直至$V_D=V_G$。继续增大漏极电压($V_D>V_G$),由于靠近漏端的栅氧化层中的电场改变方向,栅极电流将保持不变或者略微减小。图10.17描述了不同的V_G和V_D下沟道和氧化层中的电场。当V_D小于V_G时(见图10.17(a)),沟道漏端(此处y方向的电场最大)氧化层中的电场方向从栅极指向沟道($\mathscr{E}_{ox}>0$),这对栅极收集注入的电子是有利的。

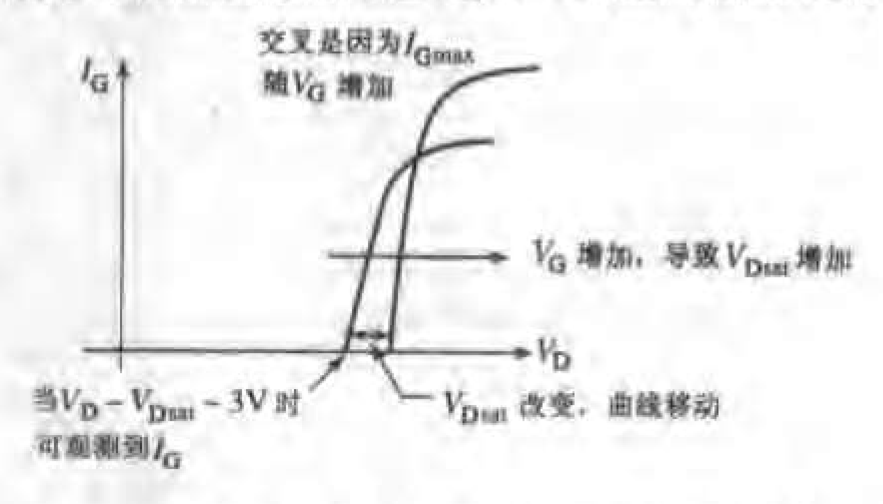


图10.16 栅极电流I_G与漏极电压V_D的关系曲线。I_G随V_D的增大先迅速增加,然后达到饱和

V_D的增大使漏极电流增加，具有足够能量克服 $Si\text{-}SiO_2$ 界面势垒的电子数目也随之增大。因为栅极电流受限于从沟道电场获得足够能量的电子的数目，这种情形下的栅极电流被称做沟道电场限制栅极电流。当 $V_D = V_G$时(见图 10.17(b))，沟道漏端氧化层中的电场为 0。当 $V_D > V_G$时(见图 10.17(c))，尽管电子有足够的能量越过势垒，漏端氧化层电场方向改变为阻碍收集热电子($\mathscr{E}_{ox} < 0$)，因此栅极电流饱和或者减小，如图 10.16 所示，这个范围内的栅极电流称做氧化层电场限制栅极电流。

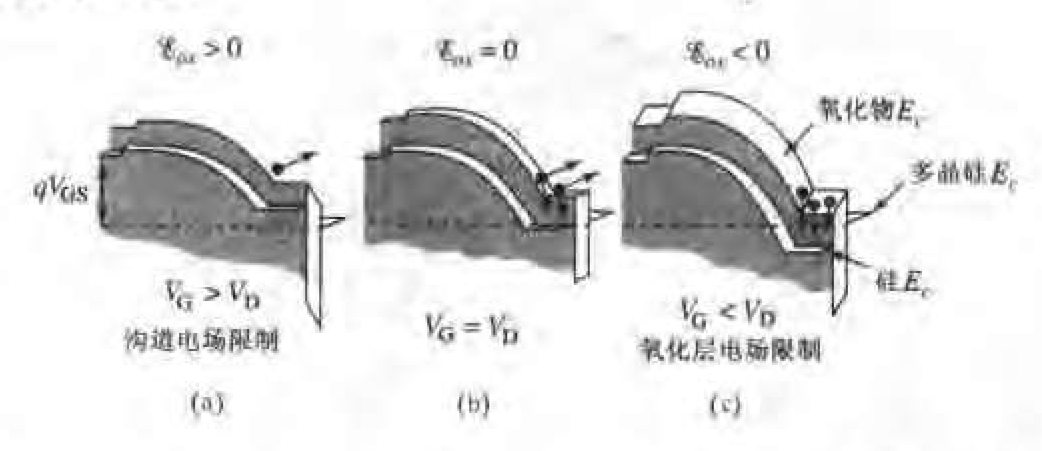


图 10.17 三种情况下漏端热电子注入的栅电场和沟道电场示意图

栅极电压较高时，$\mathscr{E}_m$ 也会下降，这将减小栅极电流，如图 10.14 所示。这种栅极电流减小的机制类似于前面讨论的衬底电流 I_{sub}随 V_G的增大而减小。

例题 MOSFET 的退化时间

已知短沟道 MOSFET，$x_{ox} = 30\text{nm}$，热电子瞬间注入产生 1nA 的栅极电流，假设发生注入的区域的面积为 $200\text{nm} \times 10\mu\text{m}$，并假设有比例为 10^{-6}的注入电子在距离 $Si\text{-}SiO_2$ 界面平均 $0.1x_{ox}$处被俘获，计算在这个栅极电流下(在上述条件下)，使注入的区域的阈值电压改变 100mV 所需要的时间。

解：1nA 的栅极电流，在注入区域的电流密度是

$$10^{-9}/2 \times 10^{-8} = 5 \times 10^{-2}\ \text{A cm}^{-2}$$

因为 10^6 个电子中有 1 个电子被俘获，SiO_2 俘获电子的速率是 $5 \times 10^{-2} \times 10^{-6} = 5 \times 10^{-8}\text{C s}^{-1}\text{cm}^{-2}$。根据等式(8.5.4)，氧化层俘获的电荷将使平带电压漂移，从而改变阈值电压(等式(8.3.18))。假设这些电荷集中在一个薄层，利用等式(8.5.3)，平带电压的改变 ΔV_{FB}为

$$\Delta V_{FB} = \Delta V_T = \left(\frac{1}{C_{ox}}\right)\left(\frac{0.9x_{ox}}{x_{ox}}\right)\Delta Q_{ot}$$

或

$$\Delta Q_{ot} = \frac{C_{ox}\Delta V_T}{0.9} = \frac{1.15 \times 10^{-7} \times 0.1}{0.9} = 1.28 \times 10^{-8}\ \text{C cm}^{-2}$$

俘获这些电荷需要的时间为

$$t = \frac{\Delta Q_{ot}}{J_{ot}} = \frac{1.28 \times 10^{-8}}{5 \times 10^{-8}} = 0.256\ \text{s}$$

10.3.2　低栅压载流子注入

前面介绍的幸运电子模型解释了高栅压和高漏压下载流子注入引起的栅极电流，这种条件下，栅极电流的问题最为严重。这个模型考虑的是载流子在沟道中直接被电场加速而形成沟道热电子（CHE），并且与碰撞电离过程无关。而在低栅压下，其他导致栅极电流的机制将变得重要。图 10.18 给出了在较大的偏压范围内栅极电流与栅极电压的关系曲线[23]。图中的数据表明，除了在高栅压区有一个由幸运电子模型预言的栅极电流最大值外，在中等栅电压区也有一个栅极电流最大值，在非常低的栅极电压下，栅极电流还有可能反向，这表明有空穴注入到栅氧化层。后面这两个电流成分是由碰撞电离导致的。因为碰撞电离也会造成衬底电流 I_{sub}，所以在衬底电流最大的偏压条件下，这两种栅极电流成分最为严重。由碰撞电离导致的热电子和热空穴的注入称为漏雪崩热载流子电流（D. A. H. C）

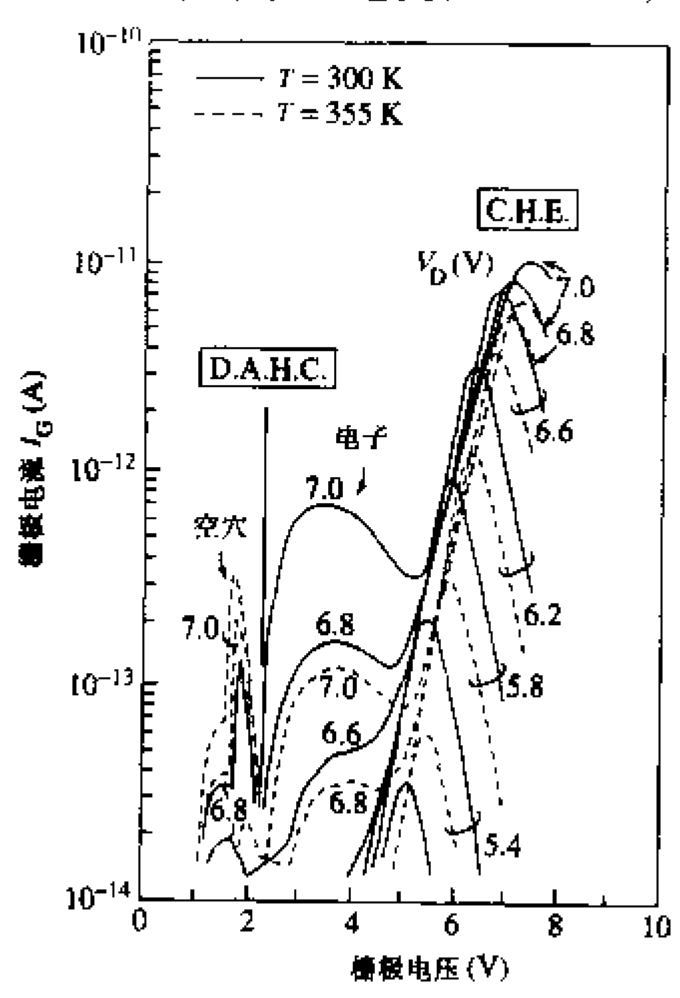


图 10.18　不同的漏极电压下测量得到的 n 沟道 MOSFET 的栅极电流与栅极电压的关系曲线。低栅压时，漏雪崩热载流子电流（D. A. H. C.）占优势。其中热空穴导致的负栅极电流在 V_G 大约小于 2V 时达到最大值；热电子引起的正栅极电流在 V_G 约为 3V 时达到最大。由幸运电子模型预言的沟道热电子（C. H. E.）在 $V_G > 4V$ 时栅极电流达到最大值[23]（© 1982 IEEE）

首先考虑低栅压下的栅极空穴电流，这些空穴是由漏区附近强场下的空间电荷区中发生的碰撞电离产生的。当栅极电压远小于漏极电压（但仍然大于阈值电压）时，漏极附近氧化层中的电场改变符号，一些有足够能量的空穴注入并穿过栅氧化层，到达栅电极，导致流出栅极的电流增大。在高漏压和低栅压情况下，这种空穴注入是引发栅极电流的主要机制。空穴电流的量级非常小，并且很难测量，尤其是当栅氧化层的厚度大于 10nm 时。在图 10.18 的数据中，当栅极电压比漏极电压小得多时（例如 $V_G = 1.2V$，$V_D = 6V$），空穴注入是栅极电流的最主要成分。当 V_G 增大到 2V（V_D 保持在 6V）时，热空穴栅极电流则减小到可以忽略。

随着栅极电压的增加,注入热电子将取代热空穴起主导作用。即使漏极附近的氧化层电场会阻碍栅电极收集注入的电子,但这些注入电子积累的能量也足以使它们中的一部分能够到达栅电极。随着漏和栅之间电场的减小(因为栅极电压增加,而漏极电压保持较高的值不变),栅极电流中的热电子成分变得尤为重要,这是因为热电子在被反方向的氧化层电场强制返回衬底之前,注入到氧化层并能抵达栅电极的概率增加了。因此,在中等栅压区,栅极电流最初的增加是由于碰撞电离而产生的电子,但是,栅极电压的继续增大将减小漏区附近沟道中的横向电场,所以,碰撞电离产生的热电子将不再对栅极电流起重要作用。由于 MOSFET 的栅极电流通常都比较小(低于 pA 量级,如图 10.18),我们给出的漏雪崩热载流子注入的理论描述及其实验验证并不全面。在更高的栅极电压下,越来越多的研究已经证实了由沟道幸运电子导致的热电子注入是栅极电流的主要成分。

10.3.3 p 沟道 MOSFET 的栅极电流

空穴的热载流子效应与电子的有所不同,并且在 p 沟道 MOSFET 中没有观察到可测量的沟道热空穴注入。产生这一现象的主要原因有两点:空穴注入到氧化层需要克服更高的势垒(见图 8.1 和图 8.2);空穴的平均自由程比电子的短。p 沟道 MOSFET 中决定栅极电流的主要因素也是电子向氧化层的注入,这与 n 沟道 MOSFET 相同。注入的电子由碰撞电离产生,并被沟道电场加速,注入的机制类似图 10.18 所示的 n 沟道 MOSFET 低栅压下的空穴注入。p 沟道 MOSFET 的栅极电流甚至可能比 n 沟道 MOSFET 大。比较 p 沟道 MOSFET 的衬底电流 I_{sub} 和栅极电流 I_G,我们发现 I_{sub} 和 I_G 都在低栅压下达到最大值[24]。

10.4 器件退化

10.4.1 n 沟道 MOSFET 的退化机制

热载流子效应导致器件的退化是随时间积累的。例如,图 10.19 给出了 MOSFET 输出特性退化的测试曲线,这个退化可以用 MOSFET 阈值电压的增加来解释。阈值电压增加将导致 MOSFET 的漏极电流下降和开关速度降低。用具有这样的退化特性的器件组成的电路最终可能会在规定的速度下停止工作,从而导致系统失效。

热载流子引起的栅极电流通过两种机制导致这种退化:(1)氧化层俘获电荷;(2)产生界面态。这两种机制都会造成负电荷在栅氧化层中和/或 Si-SiO_2 界面处积累。只要电路条件会产生热载流子注入,电荷就会不断积累,其中主要在 MOSFET 的漏端附近。这种电荷积累对被观测器件的性能影响较弱,这是因为沟道漏端的阈值电压的大小没有源端那么重要。但是,如果将器件的源极和漏极互换,退化区域将出现在源端附近,而该处的阈值电压是非常重要的。尽管评定哪种改变阈值电压的机制更为重要比较困难,但实验结果表明产生界面态这一机制更为重要。如果氧化层俘获电荷起主要作用,那么在导致栅极电流较高的偏置条件下,MOSFET 的退化应该是最严重的。但是实验发现,V_G 近似等于 $V_D/2$ 时,退化率最高,而这种偏压条件将引起衬底电流最大,而不是栅极电流最大。退化与电压的这种关系,说明退化主要是由产生界面态引起的(参考 8.5 节)[26]。产生界面态的确切机制现在仍是一个研究课题,但人们已提出了一个合理的模型,即界面陷阱态的形成是由于热电子碰撞传递了足够的能量使化

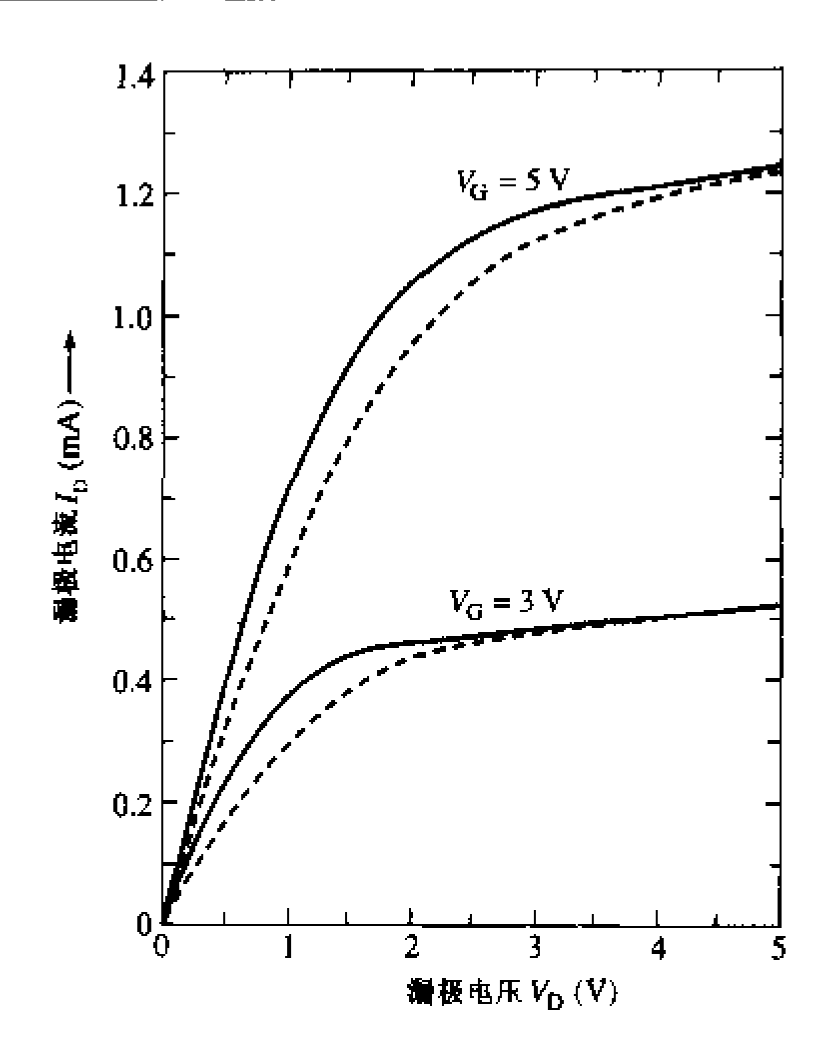


图10.19 n沟道MOSFET输出特性曲线，实线为施加热载流子应力之前，虚线为施加热载流子应力之后（$W/L = 8\mu m/2\mu m$，$x_{ox} = 28nm$，$\Delta V_T = 220mV$）[19]（© 1989 Academic Press）

学键被打破（可能是硅原子与氢原子之间的化学键）所致。

界面处硅原子之间的许多键都被单晶硅与非晶氧化层相连的过渡区中无序的晶格结构削弱或破坏了。IC工艺最后的流程之一是晶片在氢环境中退火，氢通常能与那些没有与其他硅原子或氧原子束缚的硅原子成键。这些Si-H键比Si-Si键更弱，更容易被打破，所以当热载流子穿过这个区域时，它就成了产生界面态电荷积累的区域。

10.4.2 n沟道MOSFET退化的表征

MOSFET退化的严重程度通常用阈值电压的改变量（ΔV_T），或者线性区（低V_D和高V_G，例如$V_D = 0.1V$，$V_G = V_{DD}$）漏极电流变化的百分比（$\Delta I_D/I_D$）来表征。在某些情况下，跨导（g_m）或亚阈值斜率（S）也被用来表征MOSFET的退化。

下面将讨论如何确定MOSFET是否可以在要求的电路条件下可靠地工作。常用的标准是在给定的工作时间内，热电子损害造成的器件退化在允许范围内，通常采用要求MOSFET阈值电压的最大改变量为10mV，这相当于MOSFET的漏极电流大约退化3%。这是电路可以忍受并且仍能正常运行的最大改变量。

为了测量器件的退化特性，首先需要确定一个反映应力退化的量，并开发一种测量方法作为"监测器"来测量应力积累。衬底电流I_{sub}可以作为一个被监测的量。测量期间被监测量的数值以及监测时间代表了MOSFET中应力积累的程度，而其他的量，例如阈值电压的改变量则用于判断器件是否"失效"。

应力测量的偏置条件可以与电路要求的工作条件相同，也可以让器件工作在"过应力"状

态下,例如使用更高的漏极电压 V_{DD}。栅极电压 V_G 则被调整到使衬底电流 I_{sub} 最大,这样退化测量就在最坏的情形下进行。然后,器件将在这种偏置条件下持续工作一段时间,同时监测器件的 I_{sub}。测量会被定期中断,以测量那些被选择用来指示器件失效的参数(例如 ΔV_t 或 ΔI_D)。应力测量将一直延续至达到失效标准,这时的测量时间 τ 称为"热载流子器件寿命"或"MOSFET 失效时间",或者简称为"MOS 寿命"。

10.4.3 器件寿命的加速测量

通常要求器件能在正常条件下可靠地工作 10 年,而花 10 年的时间去判断器件的可靠性是不现实的,因此,需要一种方法加速应力测量,在合理的时间内测出失效时间 τ。正如刚刚指出的,我们可以在高于正常偏置的条件下进行应力测量。一般采用增大漏极电压的方法来加速寿命测量,这时的监测量(例如 I_{sub})比正常偏置的要高一些,所以 τ 更短。

失效时间 τ 和 I_{sub} 之间是幂函数关系,一般形式是[19]

$$\tau \approx K_1\left(\frac{I_{sub}}{I_D}\right)^{-m} \tag{10.4.1}$$

式中 K_1 和 m 是经验参数。首先在几个不同的高偏压下测出 MOS 寿命 τ,然后利用等式(10.4.1)外推得到正常工作条件下的 τ。因为 I_D 与沟道宽度 W 成正比,所以需要测量不同沟道宽度的 MOSFET,并利用另一个能反映 τ 和 W 的函数关系来评估 MOS 工艺的可靠性。

$$\tau \approx K_2\left(\frac{I_{sub}}{W}\right)^{-m} \tag{10.4.2}$$

等式(10.4.1)和(10.4.2)中的一些参数的典型值为 $m=3$, $K_2=10^6$, $K_1=3$,其中 τ 的单位为秒,I_D/W 的单位为 mA/μm。图 10.20 是一个典型的 τ 与 I_{sub} 的关系曲线,该图表明 τ 与 I_{sub} 之间的幂函数关系是正确的。

例题 MOSFET 的失效时间

已知 MOSFET 的参数如下,计算 $V_{DD}=5\text{V}$ 时的失效时间 τ。如果要求器件能可靠地工作 10 年,V_{DD} 的最大值是多少?

器件参数:$x_{ox}=20\text{nm}$, $x_j=0.2\mu\text{m}$, $V_{Dsat}=1\text{V}$, $V_T=0.7\text{V}$,沟道长度 $L=1\mu\text{m}$。失效标准为 $\Delta I_D/I_D=10\%$,已知 $K_1=10$, $m=3$。

解:利用等式(10.1.16),得到

$$\ell = 0.22x_{ox}^{1/3}x_j^{1/2} = 0.124\ \mu m$$

由等式(10.1.15),得到

$$\mathscr{E}_m \approx \frac{V_D - V_{Dsat}}{\ell} \approx 3.23\times 10^5\,\text{V cm}^{-1}$$

利用等式(10.2.6),并取 $A_i/B_i=1.2\text{V}^{-1}$,有

$$\frac{I_{sub}}{I_D} = 1.2(V_D - V_{Dsat})\exp\left(-\frac{B}{\mathscr{E}_m}\right) = 0.025$$

最后,利用等式(10.4.1),得到

$$\tau = K_1\left(\frac{I_{sub}}{I_D}\right)^{-3} = 7.4\text{天}$$

注意,本例题中没有明确用到 L 和 V_G,这是因为给定偏置条件下的 V_{Dsat} 已经给出。

10 年的寿命要求 $\tau = 3.15 \times 10^8$s,所以需要

$$\frac{I_{sub}}{I_D} = 3.17 \times 10^{-3} \Rightarrow \mathscr{E}_m < 2.4 \times 10^5\ \text{Vcm}^{-1}$$

相应的 $V_{DD(max)} \approx 4.22$V

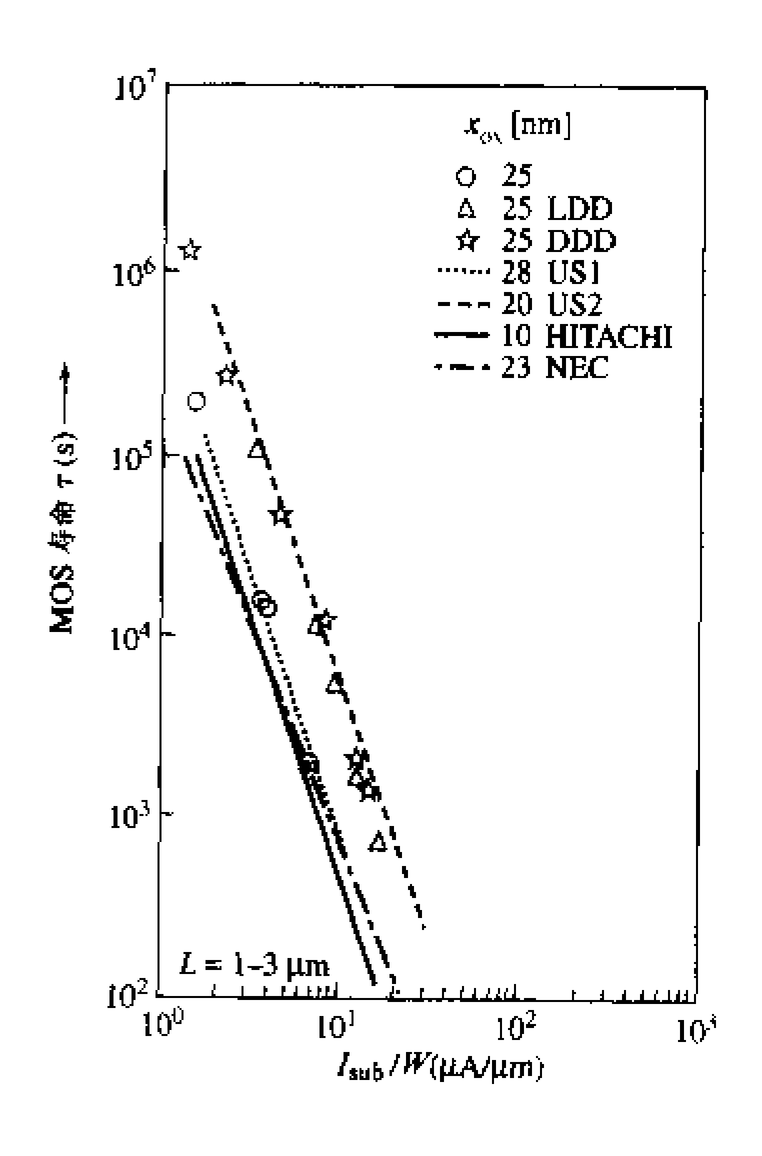


图 10.20 不同工艺下的器件寿命 τ 与衬底电流 I_{sub} 的函数关系,表明 τ 和 I_{sub} 之间存在幂函数关系[26] (© 1987 IEEE)

10.4.4 减小漏场的结构

前面的例题说明,5V 的工作电压就会导致本题考虑的 1μm CMOS 工艺的失效时间过短。降低电源电压 V_{DD} 可以增加 τ,但同时会减小可利用的栅压 V_G 的范围,因此也减小了最大驱动电流 I_{Dsat},这样就会降低电路的工作速度,影响系统性能。一种较好的解决方法是:保持电源电压不变,重新设计电路结构,使过剩漏极电压($V_D - V_{Dsat}$)降落在其他地方,而不是直接降落在速度饱和区,这样就可以减小最大电场 $\mathscr{E}_m$。这种设计不会影响到电流驱动能力,因为 I_{Dsat} 对 V_D 的依赖关系较弱。可以简单地通过在漏极增加一个串联电阻来限制沟道压降,但这种方法不是很有效,因为为了避免降低电路性能,电阻上的总压降必须非常小。一种更有效的限制沟

道中最大电场的方法是在高掺杂的漏区和沟道之间加入一个轻掺杂的缓冲区(n^-)[27]。

可以通过多种方式在 MOSFET 中加入轻掺杂的缓冲区,其中应用最广泛的是轻掺杂漏(LDD)工艺,如图 10.21 所示。形成这种结构的方法是,首先刻蚀出栅电极(通常用多晶硅制成),然后在源区和漏区注入中等剂量的 n 型杂质(浓度为 N_{d1}),并利用多晶硅的边缘来确定注入区的一个边界。接下来制造覆盖靠近栅边缘 n^- 区的侧墙隔离氧化物,然后再进行一次 n^+ 注入以形成正常的重掺杂源区和漏区。此时隔离区下方的 n^- 区仍然保持低掺杂,可用来降低不需要的过剩漏极电压[28]。

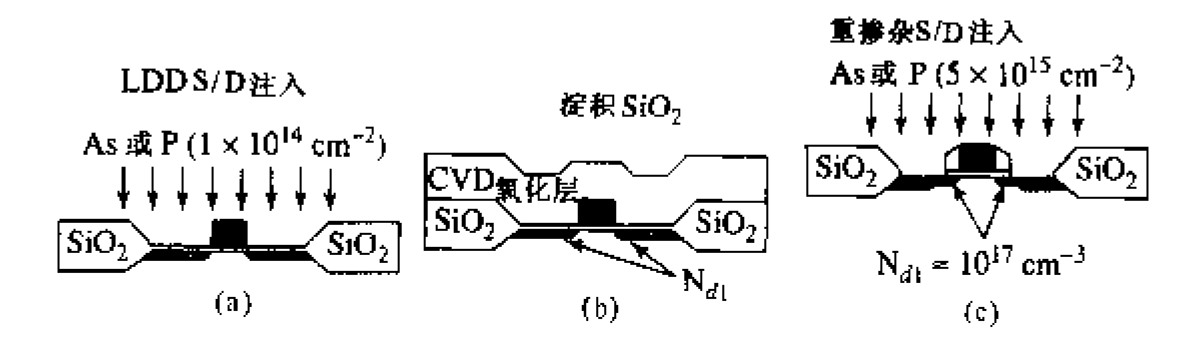


图 10.21 轻掺杂漏(LDD)结构的工艺流程:(a)注入中等剂量的 n 型杂质,(b)淀积 CVD 氧化层,(c)各向异性刻蚀 CVD 氧化层以形成"侧墙隔离",然后注入高剂量的 n 型杂质

最佳的轻掺杂漏结构要求 n^- LDD 区是完全耗尽的,形成的正的空间电荷密度等于 qN_{d1}。这部分电荷改变了速度饱和区的总电荷,等式(10.1.7)改写成

$$\frac{d\mathscr{E}(y')}{dy'} = \frac{[V(y') - V_{Dsat}]}{\ell^2} - \frac{qN_{d1}(y')}{\epsilon_s} \tag{10.4.3}$$

从等式(10.4.3)可以看出,LDD 结构减小了漏区附近沟道最大电场$\mathscr{E}_m$。从 LDD 结构和非 LDD 结构的电场分布比较图(见图 10.22)中可看出 LDD 结构对性能的改善。为了能有效地降低过剩漏极电压,控制 n^- 区的掺杂浓度是非常关键的。如果掺杂浓度过低,则串联电阻过高,限制了 I_{Dsat};如果掺杂浓度过高,LDD 区不能完全耗尽,电场的降低也就不明显。一种比较有用的设计目标是让 LDD 区的电场保持为常数,那么这种情况下沟道最大电场$\mathscr{E}_m$为

$$\mathscr{E}_m \approx \frac{V_D - V_{Dsat} - \mathscr{E}_m L_{n^-}}{\ell}$$

$$= \frac{V_D - V_{Dsat}}{\ell + L_n} \tag{10.4.4}$$

式中 L_{n^-} 是轻掺杂漏区的长度。为了制造满足上述目标,必须精确控制 LDD 区的掺杂浓度。

另一种形成 n^- 缓冲区的结构是双扩散漏(DDD)结构。为了形成这种结构,两种不同的杂质(磷和砷)被同时注入到漏区,其中磷为中等剂量(例如,$1\times10^{14}\,cm^{-2}$ 至 $1\times10^{15}\,cm^{-2}$),砷为高剂量(例如 $5\times10^{15}\,cm^{-2}$),如图 10.23 所示。在后序的高温退火工艺中,由于磷的扩散速度比砷快,在漏结处形成了一个浓度梯度较缓的区域(在通常的工艺条件下,源区也将包含双扩散。虽然源结不需要缓变过渡区来限制电场,但阻碍第二种杂质注入源区会增加工艺的复杂度(也就是说,需要增加掩膜版,因此提高了成本),这是不合理的)。虽然形成双扩散漏区的工艺比较简单,但很难控制 n^- 区的长度,并且增加了结深,这对短沟器件是不利的。所以现代工艺中并没有广泛使用 DDD 结构

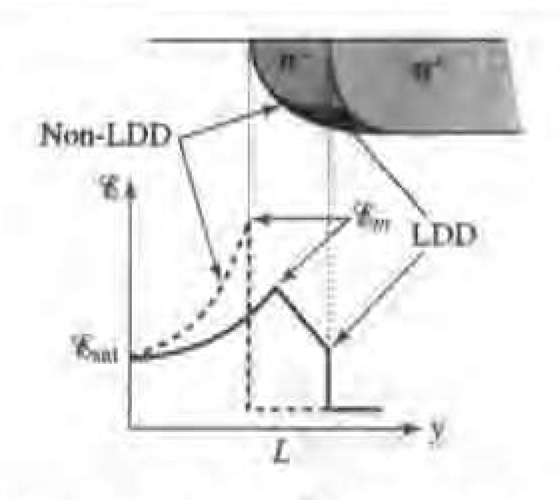


图10.22　n沟道MOSFET的电场分布，实线为LDD的结构，虚线为非LDD结构



图10.23　通过注入两种扩散速度不同的杂质得到的双扩散漏(DDD)结构

10.4.5　p沟道MOSFET的退化

p沟道MOSFET的热载流子效应没有n沟道MOSFET严重，这是因为空穴的碰撞电离率通常比电子的碰撞电离率低1－2个数量级。p沟道MOSFET一般不需要轻掺杂漏结构(尽管在CMOS电路中为了保持制造工艺与n沟道MOSFET的兼容可能会采用这种结构)。另外，与n沟道MOSFET相反，经过一段时间的工作后，p沟道MOSFET的漏极电流经常是升高而不是降低。

p沟道MOSFET退化机制可以这样定性解释[29]：在强沟道电场下，空穴被加速到足以产生碰撞电离，从而产生电子-空穴对。如果产生的一部分电子从沟道电场中获得足够的能量，并被散射到指向表面的方向，它们就可以注入到栅氧化层中，其中一部分电子在栅氧化层中被俘获并积累，尤其是在漏端附近(见10.3节)。

p沟道MOSFET中碰撞产生并被俘获的热电子主要积累在漏端附近的氧化层中，这些负电荷最终将导致漏端附近的n型硅表面区反型，相当于延伸了p^+漏区，并有效地缩短了沟道长度，有效沟道长度缩短增大了漏极电流。这种漏极电流增加效应比漏极电流减小对电路性能的破坏要小。但是，沟道长度的缩短会导致更高的泄漏电流或者导致穿通(9.2节)。沟道长度减小还会增加最大沟道电场，使p沟道MOSFET的平均失效时间对沟道长度的缩小更为敏感。

10.5　器件：MOS不挥发存储器结构

迄今为止，我们的讨论集中在MOSFET中由热电子效应引起的一些问题，主要是器件性能的退化。但是，一些重要的商用器件却是依靠与热载流子效应有关的机制来工作的，其中最主要的是不挥发存储器，即没有外加偏置时也能保存存储信息的器件。不挥发存储器的种类很多，包括将电荷储存在电绝缘电极上的浮栅存储器，通过在界面附近俘获电荷进行操作的金属-氮化物-氧化物-硅(MNOS)器件，以及铁电随机存储器(FRAM)。

我们主要讨论浮栅存储器结构，其工作原理与本章前面讨论的强场效应有关。这类存储器可分为两大类：可以电学写入，紫外线擦除的可擦除可编程只读存储器(EPROM或UV EPROM)；可以电编程和电擦除的电擦除可编程只读存储器(EEPROM或E^2PROM)。在常规的E^2PROM中，任何单元(位)都可以被随机地读或写。FLASH E^2PROM的存储单元采用与

E^2PROM 相同的基本结构，但是其阵列结构与常规的 E^2PROM 不同。为了使 FLASH E^2PROM 达到高集成度，一个单独的存储单元是不能直接进行存取的。FLASH 存储器的擦除或清零是以器件组（称为"块"）为单位进行的，然后通过将某些位的状态改变至与擦除状态相反，将信息顺序写入块中的某些位。

浮栅存储器的结构如图 10.24(a) 所示。这种器件实际上是一个有两层栅（控制栅和浮栅）的 n 沟道 MOSFET。控制栅连接外部电路，浮栅没有外部连接。由第 8 章知，经过足够的时间后，因为电路中的通路或寄生通路的电荷传输，一个 MOS 系统通常能达到热平衡要求的电荷状态。浮栅最终达到的电荷状态取决于系统中各材料的功函数。然而，电子从浮栅（通常是多晶硅）的导带运动到周围氧化物的导带需要克服较高的势垒(3.1eV)，并且氧化物的电阻很高，所以储存在浮栅中电子的泄漏非常慢。这种器件可以在非热平衡的亚稳态状态下保持数十年，因此它可被用做存储元件，在使用期间保存电荷。图 10.24 中的浮栅元件包括两个绝缘薄层，一个是衬底沟道上方的热生长形成的氧化层（厚度为 d_2）；另一个是高质量的淀积氧化层（厚度为 d_1，通常比 d_2 大），图 10.24 中覆盖在其上方的控制栅极通常是淀积的多晶硅。

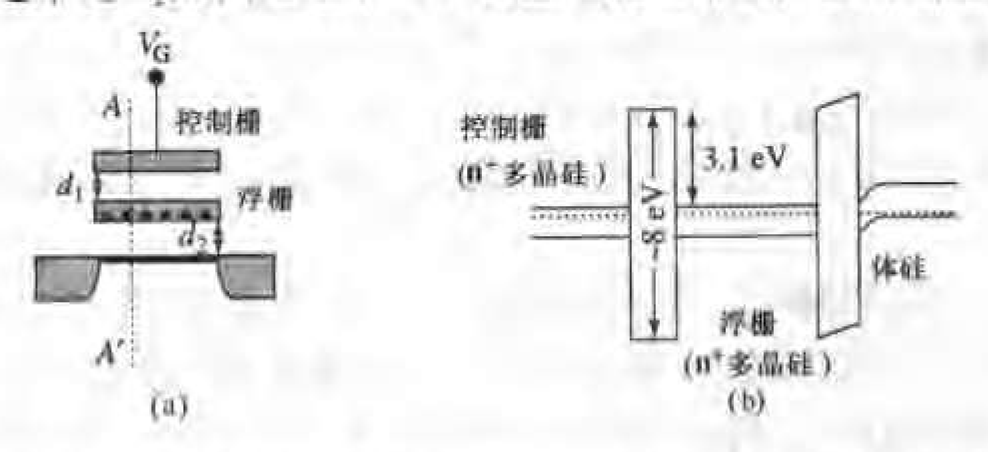


图 10.24 (a)浮栅存储器单元结构；(b)存储元件 AA' 剖面的能带图

表征栅电极感应衬底中可动电荷能力的阈值电压与存储在浮栅中的电荷总量 $-|Q_{fg}|$ 有关。阈值电压表达式为：

$$V_T = V_{FB} + 2|\phi_p| + \frac{|Q_d|}{\epsilon_{ox}}(d_1 + d_2) + \frac{|Q_{fg}|}{\epsilon_{ox}}d_1 \tag{10.5.1}$$

图 10.24(a) 的结构沿 AA' 剖面的能带图如图 10.24(b) 所示。图 10.25 表明浮栅存储的负电荷 $-|Q_{fg}|$ 增加了存储器的阈值电压，并因此使测量出来的 $I_D \sim V_G$ 曲线向正栅压方向平移。该图还给出了工艺起伏导致的 V_T 值的分布，但图中没有给出相应的 I_D 随 V_T 的变化，以便使图表示得更清楚。图 10.25 的两种存储器状态取决于存储单元的浮栅上是否存储了电荷 Q_{fg}。因此，只要浮栅没有被写入足够的负电荷 $-|Q_{fg}|$，当栅极电压处于两种状态的阈值电压之间时，就会引起单元导电。因为存在两层交叠的电极，使得浮栅存储结构不同于基本的 MOSFET，其漏极电流的计算非常复杂。定量估算漏极电流不仅需要考虑 Q_{fg} 使阈值电压的改变，还需要在基本的 MOSFET 方程中考虑等式 10.5.1 对阈值电压的修正。为了正确地分析器件，首先必须用图 10.26 所示的电容网络计算出浮栅上的电压，然后将这个电压作为栅氧化层厚度为 d_2 的 n 沟道 MOSFET 的栅极电压。用于计算漏极电流的阈值电压等于使硅中感应出沟道时的浮栅电压。定量计算漏极电流并不是毫无用处，但是对于一个不挥发存储器件，我们主要需要知道的是这个器件导通还是不导通，与之相比，准确计算漏极电流的值就显得没有那么重要了。

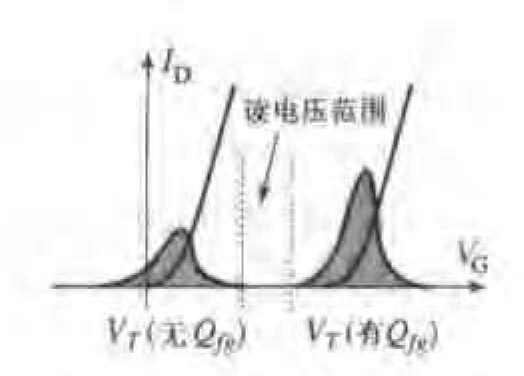


图 10.25 浮栅存储单元的 I_D-V_G 曲线。阴影区域代表了浮栅无存储电荷($Q_{fg}=0$)和浮栅存储负电荷 $-|Q_{fg}|$ 两种情况下阈值电压的改变

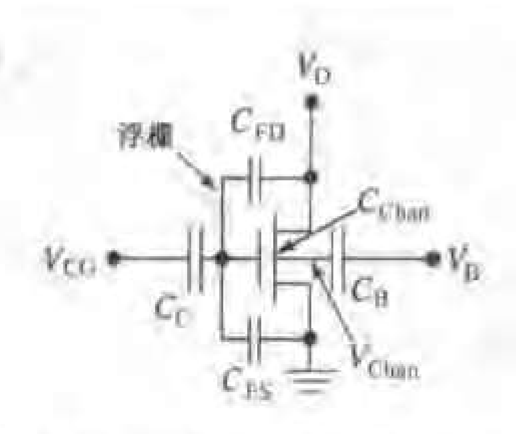


图 10.26 计算浮栅电压的电容模型

10.5.1 浮栅存储单元的编程

浮栅存储单元的读（即确定浮栅存储单元中的存储信息）比较简单直接，而写（将电子注入或移出浮栅）相对读取或访问更为复杂。写操作中，向浮栅注入电子的过程通常称为编程，消除浮栅上过剩电子的过程称为擦除。浮栅存储器的基本操作是读、写和擦除。与读过程不同，编程和擦除一般都需要比电源电压 V_{DD} 更高的电压，这些高于 V_{DD} 的电压通常是由片内的电荷泵电路产生的[30]。

10.3 节描述的沟道热电子注入是浮栅存储阵列（比如 FLASH E^2PROM）编程操作时采用的最重要的机制。图 10.27 给出了沟道漏端附近热电子注入的过程，编程电流由等式（10.3.1）给出的幸运电子模型描述。设计一个在漏端注入沟道热电子（CHE）的单元需要采取一些折衷措施：为了能有效地编程，希望栅极电流较大，所以应缩短沟道长度来提高横向电场；但是需要保证存储阵列中相邻的、没有被选中的列的单元在编程电压下不会穿通，因此，沟道长度有一个最小值。另外还有几个因素限制了浮栅存储器件的效率。首先，热电子只是沟道电流的一小部分，大部分沟道电子没有足够的能量越过界面处的势垒，所以编程过程中大部分功率都被浪费了。另外，正如 10.3 节讨论的，最佳编程电压也存在相互矛盾的要求：为了获得产生热电子所需要的强沟道电场，浮栅应该处于较低的电压；但是当浮栅电压很低时，电场的方向主要是由沟道指向浮栅，这又阻挡了热电子注入浮栅。所以为了能有效地收集热电子，需要较高的浮栅电压。由于这些相互矛盾的要求，注入效率（注入电子与总沟道电子之比）通常只有大约 10^{-6}。

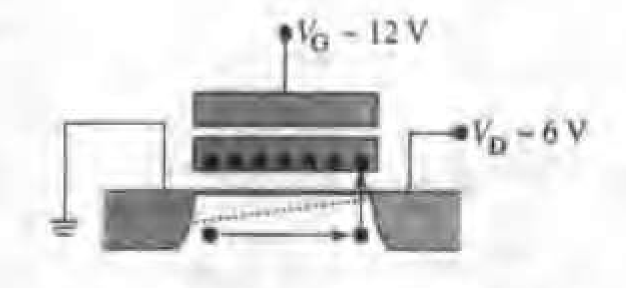
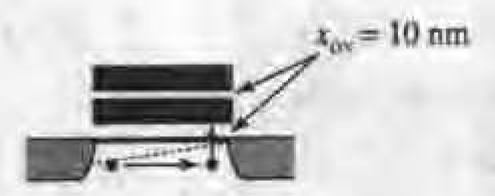


图 10.27 浮栅存储单元的热电子注入编程示意图

例题 Flash E^2PROM 的编程电流

一个 n 沟道 FLASH E^2PROM[④] 单元利用栅极热载流子电流编程。已知器件的参数如下:$V_T=0.7V$, $x_j=0.2\mu m$, $W=100\mu m$, $L=0.5\mu m$, $(V_{FB}+2|\phi_p|)=-0.2V$。为简单起见,把器件看做是 MOSFET 与一个栅电容串联,即浮栅电压由 C_{GS} 与控制栅-浮栅电容分压给出。假设浮栅中最初没有任何电荷,估算 MOSFET 加载偏置 $V_{GS}=10V$, $V_{DS}=5V$ 时的瞬时编程电流。



解:

利用等式(10.5.1)求出从控制栅测出的 V_T

$$V_T = V_{FB} + 2|\phi_p| + \frac{|Q_d|}{\epsilon_{ox}}(d_1+d_2) + \frac{|Q_{fg}|}{\epsilon_{ox}}d_1 = 0.7\text{ V}$$

由已知参数 $V_{FB}+2|\phi_p|$、d_1、d_2 和 Q_{fg} 得到

$$\frac{|Q_d|d_1}{\epsilon_{ox}} = 0.45\text{ V}$$

因此,由浮栅测出的阈值电压 V_{TF} 为

$$V_{TF} = V_{FB} + 2|\phi_p| + \frac{|Q_d|d_1}{\epsilon_{ox}} = 0.25\text{ V}$$

热电子编程时,MOSFET 被偏置在饱和区。利用 9.1 节中的模型得到,$C_{GS}=(2/3)C_{ox}$。假设该器件只是一个简单的分压器,可得到浮栅电压为

$$V_{GF} = 10\frac{C_{ox}}{C_{ox}+(2/3)C_{ox}} = 6\text{ V}$$

用 9.2 节中的方法计算漏极电流

$$\mathscr{E}_{eff} \approx \frac{V_G - V_T}{6x_{ox}} + \frac{V_T + 0.5}{3x_{ox}} \approx 1.21\times10^6\text{ V cm}^{-1}$$

$$\mu_{eff} \approx \frac{\mu_0}{1+(\mathscr{E}_{eff}/\mathscr{E}_0)^\nu} = 187\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$$

$$V_{Dsat} = \frac{(V_G - V_T)\mathscr{E}_{sat}L}{(V_D - V_T) + \mathscr{E}_{sat}L} = 2.45\text{ V}$$

$$\mathscr{E}_{sat} = \frac{2v_{sat}}{\mu_{eff}} = 8.55\times10^4\text{ V cm}^{-1}$$

$$\therefore I_{Dsat} = WC_{ox}(V_D - V_T - V_{Dsat})v_{sat} = 23\text{ mA}$$

④ 译者注:原文误为 EPROM。

还可计算出$\mathscr{E}_m$

$$\ell \approx 0.22 x_{ox}^{1/3} x_j^{1/2} \approx 0.0984\ \mu\text{m}$$

$$\mathscr{E}_m \approx \frac{V_D - V_{Dsat}}{\ell} \approx 2.59 \times 10^5\ \text{V cm}^{-1}$$

加载偏置后的瞬时充电电流由等式(10.3.2)给出

$$I_D \approx C I_{Dsat} \exp\left(\frac{\phi_B}{q\mathscr{E}_m \lambda}\right) \approx 2 \times 10^{-3} \times 23\ \text{mA} \times \exp\left(\frac{-3.32}{2.59 \times 10^5 \times 7.3 \times 10^{-7}}\right) \approx 1.1\ \text{pA}$$

一种可以显著提高注入效率的方法是采用沟道热电子源端注入。为了实现这个方法，需要在源端主浮栅旁边增加一个小的浮空注入栅，如图 10.28 所示[31]。当控制栅上加高电压时，主浮栅被控制栅容性偏置到一个高的正电压，并使大部分沟道导电良好。但是，由于几何尺寸较小，电容耦合到注入栅的电压只是略高于阈值电压，这个较低的有效偏置电压限制了注入栅下方的沟道电流，并且在注入栅靠近漏侧的边缘产生一个强电场区。注入栅下的电子在强电场中被加速，当它们到达注入栅靠近漏侧的末端并进入浮栅下方的沟道时，已经成为热电子。向浮栅提供的高的正电压正好满足收集热电子的条件，但是对注入栅下方的沟道几乎没有影响。这样，我们就解决了前面提到的矛盾。源端注入的注入效率通常约为 10^{-3}。

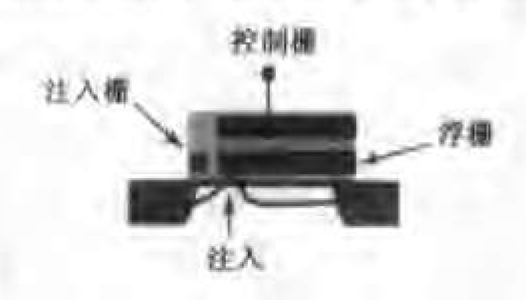


图 10.28　源端注入的结构示意图

10.5.2　浮栅存储单元的擦除

擦除指的是从浮栅中移出电子的操作。可擦除可编程 ROM 的擦除是通过紫外线激发浮栅中的电子来实现的，在这个过程中所有的器件电极都接地。相邻电子之间的排斥作用能使大部分电子从浮栅中移出。虽然这个机制很简单，但这个过程却很慢，并且会同时擦除芯片上所有的单元。另外，芯片通常需要从工作电路板转移到一个可以被照射的地方，而不是在电路板上使用电学方式直接擦除。由于这些限制，紫外线擦除 PROM 没有其他类型的可编程存储器使用广泛。

为能实现电擦除，通常需要形成一个更薄的氧化层以实现 Fowler-Nordheim 隧道效应，这种效应是大部分 E^2PROM 擦除的主要机制。Fowler-Nordheim 隧道效应是指电子向势垒顶端热激发一段距离后隧道穿通一窄势垒区进入硅的导带，如图 10.29(a)所示。图 10.29(b)指出，Fowler-Nordheim 隧道电流可以在浮栅与源极、漏极、衬底极甚至控制栅极之间流动，具体情况与偏置以及器件结构有关。例如，当加在控制栅极的电压约为 -10V，加在 p 型衬底极的电压约为 +5V，并且源极和漏极都浮空时，在浮栅与 Si-SiO_2 界面处的衬底空穴积累层之间会形成垂直的电场，电子将从浮栅隧穿到 p 型衬底。擦除时采用负的控制栅极电压可以使衬底需要的偏置电压降低到小于 V_{DD}，于是衬底可以被连接到电源而不是连接到电荷泵电路的输

出。这一点是非常重要的,因为隧穿电子电流可能高达1nA/单元,当芯片上的单元超过10^6个时,电流将达到mA量级,而片内电荷泵电路产生这个量级的电流是很困难的。另一方面,实际上没有电流流过控制栅极,因此连接到控制栅的电荷泵电路可以输出很高的电压。

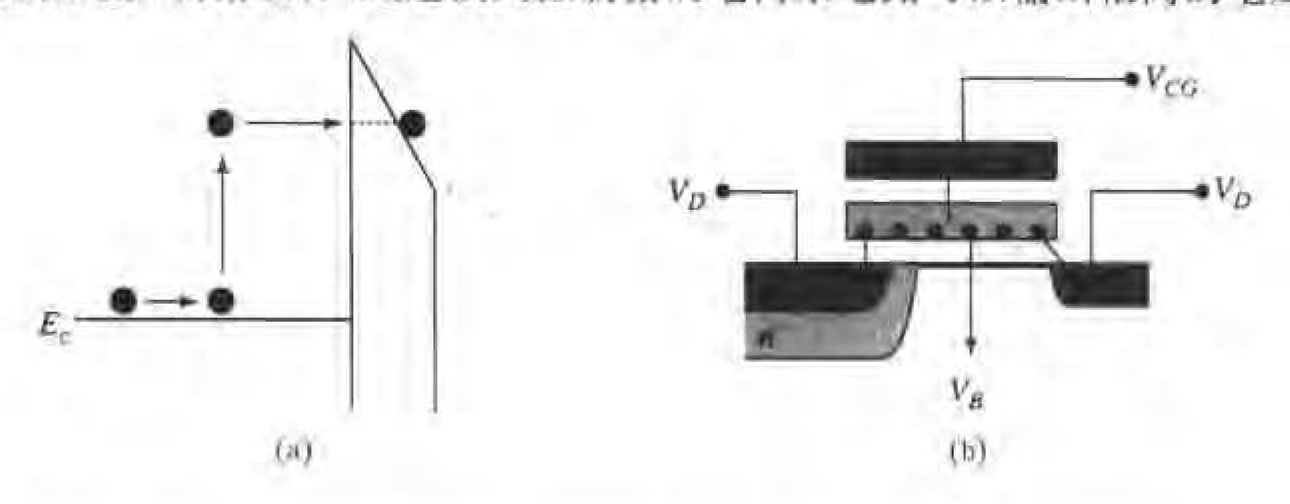


图10.29　(a) Fowler-Nordheim隧道效应;(b)擦除存储器单元的Fowler-Nordheim隧道效应的不同的可能的路径

另一种将浮栅电子擦除到源区的擦除机制也被广泛采用。在这种情况下,源极偏压大约为+5V,控制栅极加约-10V的偏压,漏极浮空。这样的偏压条件使源极附近的电场非常不均匀,如图10.30所示,n^+源扩散区和栅之间的隧穿所需的强电场耗尽了一部分n^+源扩散区,并导致这个耗尽区在栅下方向源端弯曲。这个耗尽区一方面削弱了耗尽区附近栅氧化层电场的影响,另一方面,硅中这个耗尽区的强电场会导致带带隧穿效应(BBT),如图10.30的插图所示。当穿过反偏结的电场足够高时,n^+区的价带电子能够隧穿到p型区导带中的空位,这样就发生了带带隧穿。留在n^+区的空穴将被电场加速,对器件的性能和稳定性存在潜在的影响。为了减小带带隧穿效应的影响,通常在n^+源区和p型衬底之间加一个n型区,如图10.29(b)所示。这个n型区减弱了衬底和源区之间的电场,从而相应地减弱了带带隧穿效应。

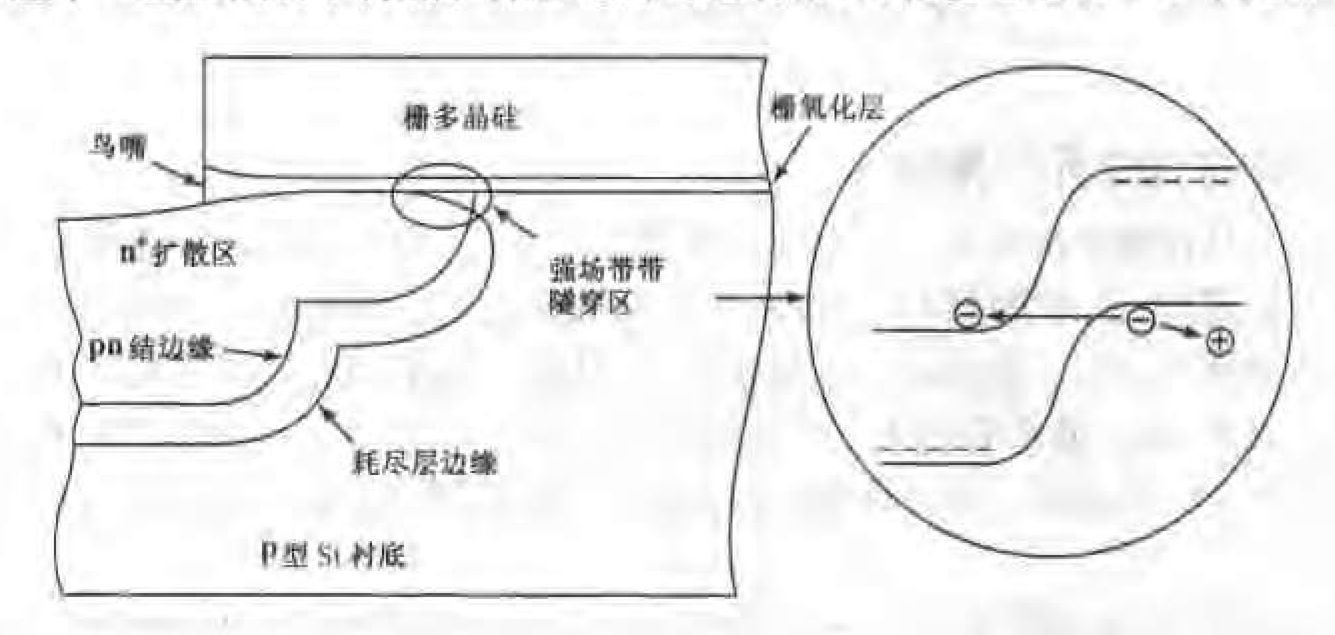


图10.30　从延伸到栅极下方的源扩散区擦除的横截面示意图[32]
(© 2000, John Wiley and Sons,授权再版)

10.5.3　浮栅存储阵列

实用的不挥发存储单元必须被规则地组成阵列来存储数字信息。随着MOS技术的提高,已经设计出越来越多的满足不同需要的不挥发存储单元,这方面的内容本文不做详细讨论。

但是,我们还是有必要举例说明不挥发存储单元如何构成阵列。考虑图10.31所示的"T"存储单元,图10.32是T-单元组成的不挥发存储阵列的排列结构图,其中横向的T-单元的控制栅极连线连接到字线,连接漏极的纵线形成位线。相邻单元的位线接触孔是共用的,所有单元的源极连接在一起。当位线和字线同时被选中时,就可以对这个单元进行编程。

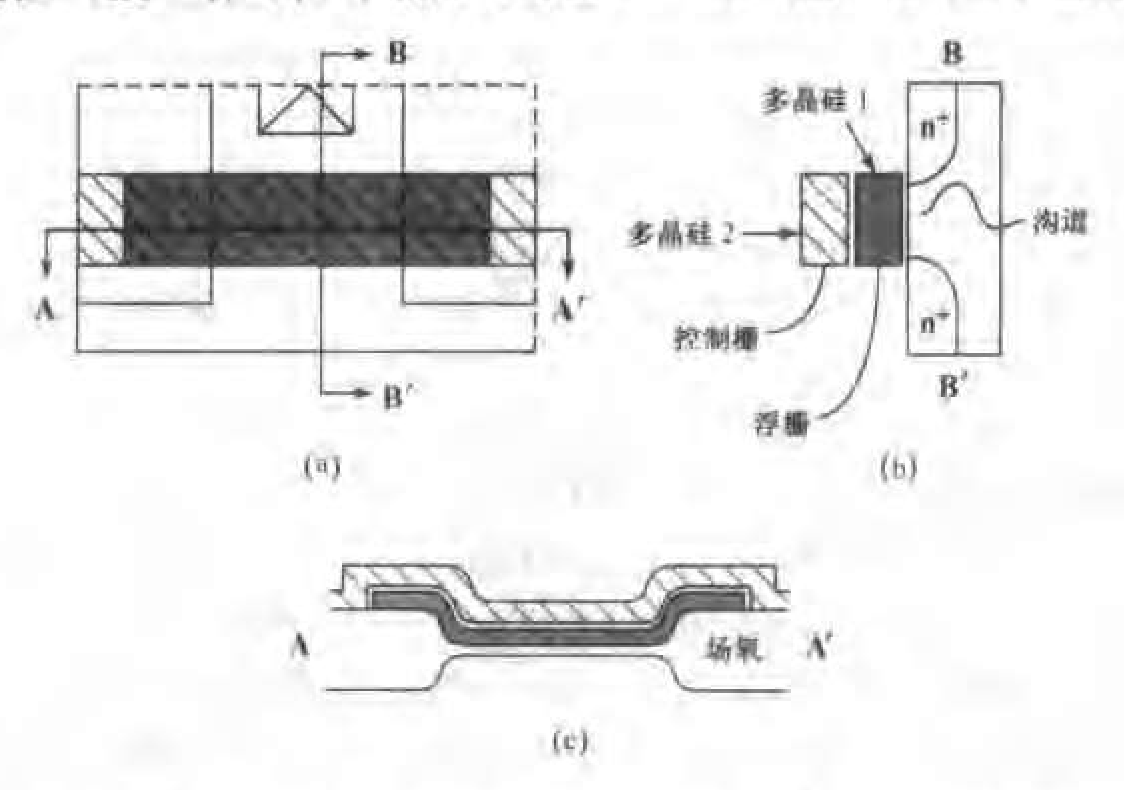


图10.31 浮栅"T"型存储单元的(a)俯视图和(b)(c)横截面图[32]
(© 2000, John Wiley and Sons,授权再版)

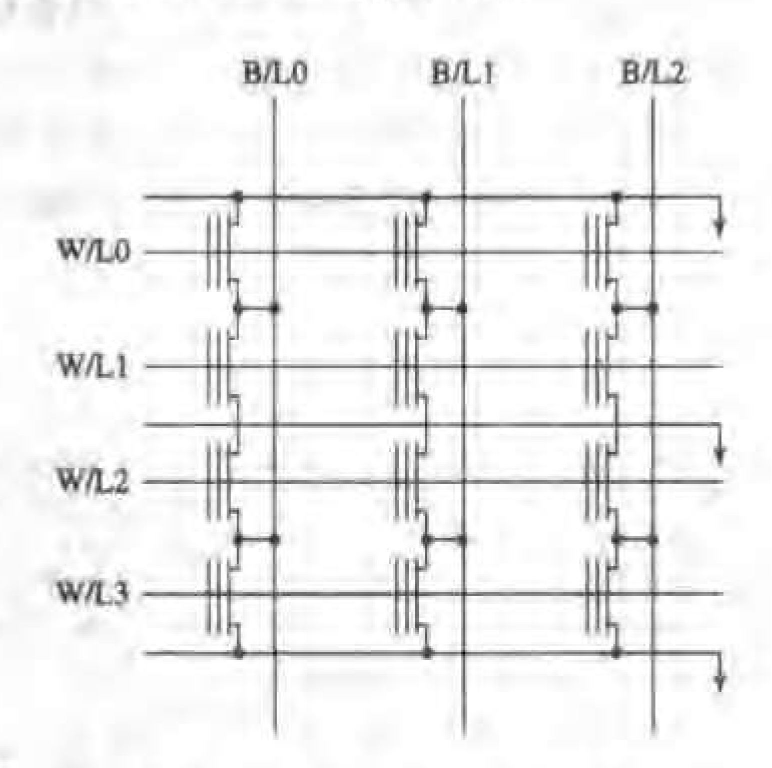


图10.32 T-单元阵列的部分结构图[32]
(© 2000, John Wiley and Sons,授权再版)

小结

亚微米MOSFET中的强场效应导致了不可忽略的衬底电流和栅极电流以及其他一些在长沟器件中并不重要的效应。为了得到短沟MOSFET中沟道区的电场分布,可以用准二维近似求解Poisson方程。准二维分析表明沟道中的电场在漏端附近迅速增加(几乎是指数增加)

到最大值,因而这个区域将发生很多重要的物理效应。迅速变化的电场导致了源漏电阻$(\partial I_D/\partial V_D)^{-1}$随着漏极电流的增加而减小。

在短的 n 沟道 MOSFET 中,输运漏极电流的电子有时候会被漏端附近的强电场"加热"至可以发生碰撞电离。电离产生的空穴在耗尽区中加速并通过耗尽区,将导致衬底电流。该电流可能使从源衬结下方的 p 型区到衬底电极之间的欧姆压降达到源-衬 np 结的导通电压,这种情况下就会使 MOSFET 失效。衬底电流的大小强烈地依赖于沟道中最大电场$\mathscr{E}_m$,而$\mathscr{E}_m$是栅氧化层厚度和漏结结深的函数。为了减小衬底电流和提高击穿电压,需要提高氧化层厚度,增大漏结结深,但这与第 9 章讨论的短沟器件的要求相互矛盾:增大结深缩短了有效沟道长度,导致更多的电荷共享,同时使器件更容易发生次表面穿通。所有这些效应都将使短沟 MOSFET 的性能下降。

漏区附近的电子具有足够高的能量,使它们能够克服沟道-氧化层势垒,产生栅极电流。一些电子在穿过栅氧化层时被俘获,导致 MOSFET 的特性退化。n 沟道 MOSFET 退化的主要机制是漏区附近 $Si\text{-}SiO_2$界面处的界面态产生,退化最严重的偏置条件是高漏压和中等栅压,此时衬底电流最大。另一方面,p 沟道 MOSFET 最重要的退化机制是氧化层俘获电子,最坏的偏置条件是高漏压和低栅压,器件退化是由于有效沟道长度缩短而导致穿通几率增加。为提高短沟道 MOSFET 的可靠性,可通过在沟道漏区附近增加一个轻掺杂漏区来减小强场效应的影响。由于其较好的性能,LDD MOSFET 已成为短沟道设计中比较受欢迎的器件。

浮栅不挥发存储器是一类很重要的 MOSFET 器件。这种器件将电荷存储在"浮"(即没有外部连线)栅上,这些电荷(在可擦除器件中既可以被存储也可以被擦除)可以改变其下方的 MOSFET 的阈值电压。为了"读取"MOSFET 存储器的状态,栅极上的偏置电压必须高于未充电器件的阈值电压,但要低于已充电器件的阈值电压。浮栅不挥发 MOSFET 存储器的充电通常采用穿过 $Si\text{-}SiO_2$势垒的热电子注入,放电采用 Fowler-Nordheim 隧道效应。

参考文献

1. T. Toyabe et al., *IEEE Trans. Electron Devices* **ED-25**, 825–834 (1978).
2. S. Selberherr, A. Schutz, and H. W. Potzl, *IEEE Trans. Electron Devices* **ED-27**, 1540 (1980).
3. M. Pinto, R. Dutton, H. Iwai, and C. Hegarty, *Tech. Digest*—IEEE Int'l Elect. Devices Mtg. (1985), p. 288.
4. G. Baum and H. Beneking, *IEEE Trans. Electron Devices* **ED-17**, 481–482 (1972).
5. A. Popa, *IEEE Trans. Electron Devices* **ED-19**, 774–780 (1972).
6. M. Fukuma and Y. Okuto, *IEEE Trans. Electron Devices* **ED-27**, 2109–2114 (1980).
7. Y. A. El Mansy and A. R. Boothroyd, *IEEE Trans Electron Devices* **ED-24**, 254–262 (1977).
8. P. K. Ko, R. S. Muller, and C. Hu, *Tech. Digest*—IEEE Int'l Elect. Devices Mtg. (1981), p. 600.
9. P. K. Ko, "Hot-electron effects in MOSFETs," Doctoral Thesis. Dept of Electrical Engineering and Computer Science, University of California, Berkeley (1982).
10. S. Tam, P. K. Ko, C. Hu, and R. S. Muller, *IEEE Trans. Electron Devices* **ED-29**, 1740 (1982).
11. S. Tam, P. K. Ko, and C. Hu, *IEEE Trans. Electron Devices* **ED-31**, 1116 (1984).
12. F. C. Hsu, P. K. Ko, S. Tam, R. S. Muller, and C. Hu, *IEEE Trans. Electron Devices* **ED-29**, 1735 (1982).
13. F. C Hsu, *IEEE Trans. Electron Devices* **ED-30**, 571–576 (1983).
14. K. Mayaram, K. Lee, and C. Hu, *IEEE Trans. Electron Devices* **ED-34**, 1509 (1987).
15. K W. Terrill, C. Hu, and P. K. Ko, *IEEE Electron Device Lett.* **EDL-5**, 440 (1984).
16. C. G. Sodini, P. K. Ko, and S. S. Wong, *Tech. Papers*, IEEE VLSI Circuits Symp., 1987, pp. 57–61.
17. Y. A. El Mansy and D. M. Caughey, *Tech. Digest*—IEEE Int'l Elect. Devices Mtg. (1975), pp. 31–34.
18. T. Y. Chan, P. K. Ko, and C. Hu, *IEEE Electron Device Lett.* **EDL-5**, 505–507 (1984).
19. C. Hu, "Hot-Carrier Effects," in *Advanced MOS Device Physics*, N. G. Einspruch and G. Gildenblat, Eds., Vol. 18, VLSI Electronics Microstructure Science, Academic Press, New York, 1989.

20. J. H. Huang, Z. H. Liu, M. C. Jeng, P. K. Ko, and C. Hu, *Tech. Digest*—IEEE Int'l Elect. Devices Mtg. (1992), pp. 569–572.
21. E. Takeda, H. Kume, T. Toyabe, and S. Asai, *IEEE Trans. Electron Devices* **ED-29**, 611 (April 1982).
22. C. N. Berglund and R. J. Powell, *J. Appl. Phys.* **42**, 573–579 (1971).
23. E. Takeda, N. Suzuki, and T Hagiwara, *Tech. Digest*—IEEE Int'l Elect. Devices Mtg. (1982), pp. 396–399.
24. K. K. Ng and G. W. Taylor, *IEEE Trans. Electron Devices* **ED-30**, 871–876 (August 1983).
25. F. H. Hsu and K. Y. Chiu, *Tech. Digest*—IEEE Int'l Elect. Devices Mtg. (1984), pp. 96–99.
26. J. Y. Choi, P. K. Ko, and C. Hu, *Digest Tech. Papers*, IEEE Symp. VLSI Tech.1987, pp. 45–46.
27. S. Ogura et al., *IEEE Trans. Electron Dev.* **ED-27**, 1359 (1980).

习题

10.1　证明当$(V_D - V_{Dsat})/\ell \gg \mathscr{E}_{sat}$时，饱和区 MOSFET 的输出电阻表达式为

$$R_{out} = \frac{dV_D}{dI_D} = \left(L + \frac{V_G - V_T}{\mathscr{E}_{sat}}\right)\left(\frac{\mathscr{E}_m}{I_{Dsat}}\right)$$

10.2　已知 n 沟道 MOSFET，$L = 1\mu m$，$x_{ox} = 20nm$，$x_j = 0.3\mu m$，$W = 100\mu m$，在对数坐标系中绘出 g_{msat}、R_{out}、$g_{msat}R_{out}$与 I_{Dsat}之间的关系曲线，绘图范围为 $V_G - V_T > 0.1V$

10.3　在热载流子导致的击穿情况下，比值 I_{sub}/I_D通常会下降 0.05 至 0.2。本题假设为 0.05，并且与器件无关。

(a) 已知 1μm CMOS 工艺的器件和工艺参数如下：$x_{ox} = 20nm$，$x_j = 0.3\mu m$，$V_{Tn} = 0.7V$，$V_{Tp} = -0.7V$。推导由于热电子导致的 n 沟道 MOSFET 的击穿电压 V_{BD}与 L 以及 $V_G - V_T$的函数关系式，绘出 $\log(V_{BD}) \sim \log(L)$曲线，$L$ 的取值从 1μm 变化到 100μm，$V_G - V_T = 2V$。等式(10.2.6)中取 $A_i = 2\times10^6 cm^{-1}$，$B_i = 1.7\times10^6 V/cm$。

(b) 对 p 沟道 MOSFET 重复(a)，取 $A_i = 8\times10^6 cm^{-1}$，$B_i = 3.7\times10^6 V/cm$。

10.4*　已知 n 沟道 MOSFET 的 $L_{eff} = 1\mu m$，$x_j = 0.3\mu m$，$x_{ox} = 20nm$，$V_T = 0.7V$。为了避免器件退化，要求$\mathscr{E}_m < 2\times10^5 V/cm$。

(a) 在这个 $\mathscr{E}_m$ 下，I_{sub}/I_D等于多少？计算最大的电源电压 V_{DD}。假设 $V_G = 3V$，并且在这个大小的栅极电压下产生最大的衬底电流。

(b) 栅极电压等于 3V 时的击穿电压是多少？假设击穿时的 $\mathscr{E}_m = 4\times10^5 V/cm$。

10.5　假设$\mathscr{E}_{sat}L \gg V_G - V_T$，推导 $\ln(I_{sub}) \sim V_G$曲线中的峰值电流表达式，并标出其位置。为了简化计算，假设等式(10.2.6)中 V_D主要通过指数项影响衬底电流，线性项中的 V_D可以忽略，即等式(10.2.6)可以被近似成

$$I_{sub} \approx \frac{A_i}{B_i} I_{Dsat} \exp\left(-\frac{\ell B_i}{V_D - V_{Dsat}}\right)$$

10.6*　如果考虑短沟道效应和 DIBL 效应，可以证明 $V_T \approx V_{T0} - 2V_D e^{-L/\ell}$，其中 V_{T0}是长沟器件的阈值电压。已知当 MOSFET 处于开启状态时，$\mathscr{E}_m \approx (V_D - V_{Dsat})/\ell$。假设为避免热载流子效应，$\mathscr{E}_m$必须小于 $2\times10^5 V/cm$，并假设 V_T的下降必须小于 0.2V，其中 V_T下降是因为(a)L 减小 0.2μm 或(b)工艺的不稳定导致 L 减小 20%。当 $L = 2\mu m$、1.2μm、0.6μm 和 0.3μm 时，计算最大的可接受的 V_D以及相应的 ℓ 值，并计算最坏情况 $L - 0.2\mu m$ 以及 $L - 0.2L$ 时的 V_{Dsat}。为简化计算，可以使用 10.5 题中的近似方法。

10.7*　氧化层的损坏导致器件“报废”的原因之一是，氧化层俘获热电子，使氧化层电场增加到氧化层的击穿电场 10MV/cm。考虑 10.3 节关于热电子导致阈值电压漂移的例题，已知起始状态时氧化层中没有电荷，穿过氧化层的电压维持在 20V，并假设热电子注入持续发生

(a)请问氧化层发生击穿需要多长时间？

(b)绘出起始状态和击穿状态时氧化层中的电场以及电压与距离的关系曲线。

(c)其他参数保持不变,假设俘获位置集中在距离栅电极 3nm 处,重复(a)、(b)。

10.8* (a) 证明 MOSFET 栅极上可以导致氧化层击穿的电子数目只是栅极面积和氧化层介电常数的函数。

(b) 如果一个 $L=6\mu m$,$W/L=5$ 的 MOSFET 被静电场充电,在栅电场达到击穿电场之前有多少电子能被转移到栅极?

(c) 如果平均充电电流为 1pA,转移这些电子需要多长时间?

10.9 已知 LDD MOSFET 的栅与 LDD 区交叠,定性绘出 LDD 区$\mathscr{E}_y(y)$曲线。如果 LDD 结构的 n 型区被本征硅(无掺杂)替换,在同一个图中绘出此时的$\mathscr{E}_y(y)$曲线。然后再绘出结构相同,但没有本征区(除去本征区)的 MOSFET 的$\mathscr{E}_y(y)$的曲线。分别说明这些曲线最重要的特征。(从这些曲线中,应该可以观察到$\mathscr{E}_m$与掺杂浓度的关系)

10.10 如果栅与 LDD 区不交叠,重复 10.9 题。

10.11 如果可以制造任意的横向掺杂分布的 n 型区,我们希望可以让 $d\mathscr{E}_m/dy=0$,这样就不会出现 10.4 节提到的电场增加的现象。为制造这种器件,n 型区沿 y 方向应该具有怎样的掺杂分布 $N_d(y)$?(并证明你的结论是最佳的)

10.12 已知非 LDD MOSFET 的参数如下:$x_{ox}=20nm$,$x_j=0.3\mu m$,$W=50\mu m$,$L=0.5\mu m$,$V_T=0.7V$,计算源极零偏压,栅极电压 $V_G=3V$,漏极电压 $V_D=3V$ 时的$\mathscr{E}_m$和 I_{sub}。当加入一个最佳掺杂分布的 0.1μm 的 LDD 区后,$\mathscr{E}_m$和 I_{sub}分别减小多少?

10.13* 已知 MOSFET 的参数和偏置条件与 10.12 题相同,计算加入与 10.12 题相同的 LDD 区后漏极电流的改变量。假设 LDD 区的电阻为 1kΩ/□。

10.14* 考虑 10.5 节关于 FLASH 存储器的例题,估算从例题中的起始状态开始,对器件编程到阈值电压等于 1V 所需要的时间。(注意:编程电流是随时间变化的,但为了计算简单,可使用平均编程电流,即起始电流和最终电流的平均值)

10.15 在 10.14 题的编程条件下,推导编程电流与注入到浮栅上的电荷(Q_{fg})的函数关系式。求出浮栅上的总电荷(Q_{fg}),并求出最大充电电流。为简化计算,本题假设 $V_{FG}-V_T \ll \mathscr{E}_{sat}L$。

部分参考答案

以下为各章习题中标有“*”的习题参考答案。

第1章 **1.2** (*a*) $E_i - E_f = 0.35$ eV, $n = 2.1 \times 10^4$ cm^{-3}, (*b*) $E_f - E_i = 0.29$ eV, $p = 2.1 \times 10^5$ cm^{-3}. **1.4** (*a*) 2.3×10^5 cm^{-3}, (*b*) 4×10^9 cm^{-3}, (*c*) 9.5×10^{16} cm^{-3}. **1.6** 160 cm^2 V^{-1} s^{-1}. **1.8** (*a*) 207 collisions, (*b*) 162 mV. **1.10** -825.6 A cm^{-2}. **1.14** (*a*) 1.85 μm (*IR*), (*b*) 1.10 μm (*IR*), (*c*) 0.87 μm (*near IR*), (*d*) 0.14 μm (*UV*).

第2章 **2.1** (*a*) 1.36×10^{15} cm^{-3}, (*b*) 2.21×10^{15} cm^{-3}. **2.4** 1.1 hour. **2.7** 0.73 μm. **2.9** (*a*) 1.79 μm, (*b*) 1.88 hours. **2.18** 34.5 ns. **2.20** (*a*) 147 Ω per square, (*b*) an added *acceptor* density of approximately 6.5×10^{16} cm^{-3} (implant *dose* of 2.6×10^{13} cm^{-2}).

第3章 **3.1** (*a*) 4.094 eV, 4.261 eV, (*b*) 0.167 V. **3.5** (*a*) $C = 0.282$ pF, (*b*) $V_R = -0.813$ V. **3.7** (*a*) $N_{d\max} = 2.9 \times 10^{17}$ cm^{-3}, (*b*) $\rho > 0.04$ Ω cm. **3.16** (*a*) for *n*-type doping, a Schottky barrier is obtained if $N_d > 7.45 \times 10^{11}$ cm^{-3}; this is smaller than any practical IC doping level, (*b*) since $\Phi_M - X_S \approx E_g/2q$, any level of *p*-doping will result in a blocking contact.

第4章 **4.1** (*a*) $\phi_i = 0.72$ V, $x_d = 0.97$ μm, $\mathscr{E}_{\max} = 1.48 \times 10^4$ V cm^{-1} (at 0 V), (*b*) $x_d = 3.73$ μm, $\mathscr{E}_{\max} = 5.75 \times 10^4$ V cm^{-1} (at -10 V). **4.3** $\mathscr{E}_{\max} = 648$ V cm^{-1}; for abrupt junction $\mathscr{E}_{\max} = 1.53 \times 10^4$ V cm^{-1}. **4.8** (*a*) $N_{do} = 6.1 \times 10^{17}$ cm^{-3}, (*b*) $a = 1.83 \times 10^{21}$ cm^{-4}, (*c*) $\mathscr{E}_{\max} = 3.4 \times 10^4$ V cm^{-1}. **4.10** (*a i*) 0.33 eV, (*a ii*) 0.36 eV, (*b i*) 0.62 V, (*c*) total depletion possible at $V_R = 23.5$ V.

第5章 **5.1** (*a*) 7.25×10^{16} cm^{-3} s^{-1}, (*b*) 2.6×10^{11} cm^{-3} s^{-1}. **5.2** (*a*) $n = 10^{16}$ cm^{-3}, $p = 10^{13}$ cm^{-3}, (*b*) 10 ns. **5.9** $\gamma = 1.55 \times 10^{-4}$. **5.11** (*a*) 0.738 V, (*b*) $n_p(-x_p) = 1.37 \times 10^{14}$ cm^{-3}, $p_n(x_n) = 6.83 \times 10^{13}$ cm^{-3}, (*c*) $J_T = 0.472$ A cm^{-2}, (*d*) currents are equal at 1.35 μm from the physical junction (in *p*-type region).

5.21

	V_a (V)	r_d (Ω)	Capacitance (pF)
(*a*)	0.1	2.95×10^{10}	3.36
	0.5	6.14×10^3	6.68
	0.7	2.8	3580
(*b*)	0	1.38×10^{12}	3.16
	−5	5×10^{95} *	1.18
	−20	10^{346} *	0.63

* other mechanisms actually limit the resistance

第6章 **6.1** 9.9×10^{14} cm^{-3}. **6.5** (*a*) 1.346 μC cm^{-2}, (*b*) 0.84 μm. **6.8** $\phi_i = 0.91$ V, $K_A = 0.163$ pC V$^{-1/2}$, $Q_{VE}(-50\text{ V}) = -1$ pC, $Q_{VE}(0.3\text{ V}) = +0.028$ pC. **6.12** $V_{CEsat} = 0.048$ V. **6.17** (*a*) $\gamma = 0.99722$, (*b*) $\alpha_T = 0.99993$, (*c*) $\beta_F = 360$, error is 2.57%.

第7章 **7.3** $\dfrac{V_A\text{ (constant doping)}}{V_A\text{ (exponential doping)}} = 0.74.$
7.7 $\beta_F = 1.333$.
7.9 $\beta_F = \beta_o[0.75 + 0.25\sqrt{1 + 4I_F/I_K}]^{-1}$ where I_F is the forward-injected current and $I_K = qD_nA_EN_a/x_B$. **7.23** (*a*) $Q_F = 24$ pC, (*b*) $Q_F = 81.1$ pC, $Q_R = 156$ pC, (*c*) charge ratio = 9.88.

7.29

	constant doping	exponential doping
g_m	77.5 mS	77.5 mS
C_D	1.94 pF	1.61 pF
δ	2.5×10^{-4}	2.08×10^{-4}
η	1.5×10^{-4}	2.2×10^{-4}

第8章 **8.1** (*a*) $V_{FB} = -0.17$ V, (*b*) $V_{FB} = -0.90$ V. **8.12** (*a*) 1.18 V, (*b*) -2.02 V. **8.15** $C_{MOS}/C_{junction} = 22.8$. **8.16** $N_{st} = 2 \times 10^{10}$ cm^{-2}.

第9章 **9.3** (*a*) $Q_n/q = 4.17 \times 10^{12}$ cm^{-2} (*b*) $V_{GS} - V_T = 3.82$ V

9.4 (*a*) at source, $v = 1.26 \times 10^7$ cm/s
at drain, $v = \infty$
(*b*) at source, $v = 3.68 \times 10^6$ cm/s
at drain, $v = 8 \times 10^6$ cm/s

9.5 (*a*) 18.7 μA (*b*) 15.5 μA (*c*) 12.2 μA (*d*) 7.56 μA **9.7** $V_T(0) = 1$ V, $k = 50\ \mu A/V^2$, $\gamma = 0.3\ V^{1/2}$, $V_A = 20$ V

9.16

	V_{DD}(V)	$\mathscr{E}_{eff}$ (MV/cm)	μ_{eff} (cm²/V-s)	$\mathscr{E}_{sat}$ (V/cm)	V_{Dsat} (V)	I_{Dsat} (mA)	g_{msat} (mA/V)
$L = 0.5\ \mu m$							
NMOS	5	0.744	307.0	5.21×10^4	1.62	49.36	15.8
	3.3	0.556	384.6	4.16×10^4	1.16	26.52	14.8
PMOS	5	0.744	92.78	12.9×10^4	2.58	23.76	8.84
	3.3	0.556	112.0	10.7×10^4	1.75	11.7	7.56
$L = 0.02\ \mu m$							Improvement
NMOS	5	0.744	307.0	5.21×10^4	0.102	77.31	56.6%
	3.3	0.556	384.6	4.16×10^4	0.081	46.39	74.9%
PMOS	5	0.744	92.78	12.9×10^4	0.243	56.04	135.9%
	3.3	0.556	112.0	10.7×10^4	0.198	33.18	183.6%

9.21 (*a*) $V_o = 0.232$ V (*b*) $V_{TD} = -2.36$ V (*c*) 0.223 V

第10章 **10.4** (*a*) $I_{sub}/I_D = 8 \times 10^{-4}$ (*b*) $V_{DD(max)} = 4.35$ V (*c*) $V_D = 7.6$ V

10.6

	2 μm		1.2 μm		0.6 μm		0.3 μm	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
ℓ (μm)	0.4	0.36	0.25	0.24	0.12	0.14	0.043	0.081
V_D (V)	9.5	8.6	5.9	5.7	2.9	3.3	1.0	1.9

10.7 (*a*) 2.56 s (*c*) 23 s **10.8** (*b*) 2.72×10^7 electrons (*c*) 4.35 s **10.13** reduced by 1.04% **10.14** 0.9 s

附　　录

符号	物理意义	章节
$A_{C,E}$	BJT 的集电结,发射结面积	6.3
A_i	碰撞电离参数	4.4,10.2
B	磁场	1.3
B	抛物线速率常数	2.3
B	隧道击穿特征电场	4.4
B_i	碰撞电离参数	4.4,10.2
B/A	线性速率常数	2.3
BV	击穿电压	4.4
C	小信号电容	3.2
C_D	混合 Π 模型中的扩散电容	7.5
C_{FB}	平带电容	8.4
C_{GD}	本征栅-漏电容	9.1
C_{GS}	本征栅-源电容	9.1
C_{RD}	寄生栅-漏电容	9.1
C_{RS}	寄生栅-源电容	9.1
C_{ID}	漏耗尽区电容	9.1
C_{IS}	源耗尽区电容	9.1
C_d	二极管扩散电容	5.4
C_g	气流中的杂质浓度	2.5
C_j	二极管结电容	5.4
$C_{jc,je}$	集电结,发射结小信号电容	7.1
C_o	气流中氧化剂浓度	2.3
C_{ox}	氧化层电容	8.3
C_p	峰值掺杂浓度	2.5
C_s	固体表面杂质浓度	2.6
C_{sHF}	高频硅表面耗尽层电容	8.5
C_{sLF}	低频硅表面耗尽层电容	8.5
C^*	气相氧化剂平衡浓度	2.3
D_{eff}	有效扩散系数	2.5

D_{it}	界面陷阱密度(单位面积和能量)	8.5
$D_{n,p}$	电子,空穴的扩散系数	1.2
D_s	表面态密度	3.5
E_I	碰撞电离所需的最小能量	4.4
E_d	施主能级	1.1
E_a	激活能	2.6
E_f	Fermi 能级	1.1
$E_{fn,fp}$	电子,空穴的准 Fermi 能级	1.1
E_g	禁带宽度	1.1
E_i	本征 Fermi 能级	1.1
E_0	自由电子能量	3.2
E_q	表面态能量	5.2
$\mathscr{E}$	电场	1.2
$\mathscr{E}_{s0}$	表面电场	9.2
$\mathscr{E}_c$	与沟道速度对应的特征电场	1.2
$\mathscr{E}_{eff}$	MOS 沟道中 x 方向的平均电场	9.2
$\mathscr{E}_l$	极限散射临界电场	7.1
$\mathscr{E}_s$	表面电场	8.2
$\mathscr{E}_{ox}$	氧化层电场	8.2
$\mathscr{E}_m$	最大电场	10.1
$\mathscr{E}_{max}$	最大电场	3.2
$\mathscr{E}_{sat}$	速度饱和临界电场	9.2
G_A	Auger 产生率	5.2
$G_{n,p}$	电子,空穴产生率	5.1
G_{sp}	自发产生率	5.2
GN	Gummel 数	6.2
I_{0E}	BJT 的发射结饱和电流	6.4
$I_{C0,CB0}$	集电极电流(发射极开路)	6.4
I_{CE0}	BJT 的发射结饱和电流	6.4
I_D	漏极电流	9.1
I_{DD}	电源电流	9.3
I_F	场感应结电流	8.6
I_G	栅极电流	10.3
$I_{KF,KR}$	Gummel-Poon 模型中的"膝点"电流	7.7
I_M	冶金结产生电流	8.6
I_S	表面产生电流	8.6

I_{pE}	BJT 的发射区空穴电流	6.2
I_{rB}	基区复合电流	6.2
I_{Dmax}	I_{Dsat}的最大值	9.2
I_{Dsat}	MOSFET 的饱和漏极电流	4.5,9.1
I_{sub}	衬底电流	10.2
J	电流密度	1.1
J_0	二极管饱和电流密度	5.3
J_1	Kirk 效应的特征电流密度	7.1
J_s	连接电流密度(BJT)	6.1
J_g	产生电流密度	5.3
$J_{n,p}$	电子,空穴电流密度	1.2
J_r	复合电流密度	5.3
J_t	二极管总电流密度	5.3
K	MOSFET 的按比例缩小因子	表 9.1
$K_{1,2}$	失效时间参数	10.4
K_I	理想因子	9.2
L	特征长度	2.10
L	FET 的沟道长度	4.5,9.1
L_D	Debye 长度	3.4
L_{Di}	本征 Debye 长度	4.2
L_n	轻掺杂漏(LDD)区长度	10.4
$L_{p,n}$	空穴,电子的扩散长度	5.3
L'	FET 未耗尽的沟道长度	4.5
M	雪崩倍增因子	4.4
M_i	模型参数值	9.3
N_m	“方盒”近似中的注入浓度	9.1
N_c	导带中的有效态密度	1.1
$N_{d,a}$	施主,受主浓度	1.1
N_i	注入杂质浓度	9.1
N_{st}	表面态面密度	5.2
N_v	价带中的有效态密度	1.1
N'	杂质面密度(注入剂量)	2.5
Q_B	基区多子电荷	6.1
Q_B	体电荷(MOSFET)	9.1
$Q_{F,R}$	正向,反向控制电荷	7.3
$Q_{DE,DC}$	发射结,集电结存储电荷	6.3

Q_d	耗尽区电荷密度	8.2
Q_f	界面固定电荷密度	8.5
Q_{fg}	浮栅存储电荷	10.5
Q_{it}	界面陷阱电荷密度	8.5
Q_m	可动载流子电荷浓度(氧化层)	8.5
Q_n	沟道自由电子电荷	8.2
Q_{nB}	存储在基区的少子电荷(npn BJT)	7.3
Q_{ot}	氧化层陷阱电荷密度	8.5
Q_{ox}	氧化层电荷	8.5
Q_{pE}	存储在发射区的少子电荷(npn BJT)	7.3
$Q_{p,n}$	少子存储电荷(二极管)	5.4
Q_s	半导体中的空间电荷	3.2
R_A	Auger 复合率	5.2
R_B	基区分布电阻	7.2
R_{BD}	源结耗尽区边缘与衬底电极之间的寄生电阻	9.1
R_{BS}	漏结耗尽区边缘与衬底电极之间的寄生电阻	9.1
R_D	漏区电阻(MOSFET)	9.1
R_H	霍尔系数	1.3
R_S	源区电阻(MOSFET)	9.1
R_W	阱电阻(CMOS)	9.3
R_X	衬底电阻(CMOS)	9.3
R_d	淀积速率	2.6
R_p	射程	2.5
R_{sp}	自发复合率	5.2
R_{out}	输出电阻	10.1
S	亚阈值斜率	9.1
T	温度	1.1
T_{tr}	渡越时间	9.1
U	净复合率	5.2
U_A	净 Auger 复合率	5.2
U_s	净表面复合率	5.2
$V_{A,B}$	Early 电压(正向,反向偏置)	7.1
V_B	MOSFET 的体(或衬底)电压	9.1
V_{BD}	击穿电压	9.4
V_C	MOSFET 的沟道电压	8.3

V_D	MOSFET 的漏极电压	9.1
V_{FB}	平带电压	8.1
V_G	MOSFET 的栅极电压	9.1
V_H	霍尔电压	1.3
V_S	MOSFET 的源极电压	9.1
V_T	MOSFET 的阈值电压	8.3
V_{Tn}	阈值电压(n 沟道 MOSFET)	9.1
V_{Tp}	阈值电压(p 沟道 MOSFET)	9.1
V_a	外加电压	3.2
V_t	热电压(kT/q)	7.5
V_{CEsat}	BJT 的饱和集电极-发射极电压	6.4
V_{Dsat}	漏极饱和电压	9.1
V_{Dsat}	漏极饱和电压	4.5
W	特征宽度	2.10
Z_{in}	输入阻抗	9.4
$e_{n,p}$	跃迁几率	5.1
f_D	Fermi-Dirac 分布函数	1.1
f_T	β_F 等于 1 的频率	7.5
$g_{\square}$	方块电阻的导数	2.10
g_m	跨导	4.5
g_{msat}	饱和区跨导	4.5
$g(E)$	电子态密度	1.1
h	普朗克常数	1.1
k	Boltzmann 常数	1.1
k_s	表面反应速率常数	2.6
ℓ	MOSFET 的结构参数	10.1
m	质量	1.1
n	电子浓度	1.1
n	二极管理想因子	3.3
n	亚阈值斜率理想因子	9.1
n_i	本征自由载流子浓度	1.1
n_{ie}	有效本征载流子浓度	1.1
n'_p	过剩电子浓度(p 型区)	5.3
n_{po}	平衡态电子浓度(p 型区)	5.3
n_s, p_s	表面电子,空穴浓度	3.3
n'	过剩自由电子浓度	5.2

p	空穴浓度	1.1
p	气相分压	2.3
p	模型参数	9.3
p'_n	n 型区过剩空穴浓度	5.3
p_{no}	平衡态空穴浓度(n 型区)	5.3
p'	过剩自由空穴浓度	5.2
q	电荷量	1.1
$q\phi_B$	能量势垒	3.2
r	俘获和发射率	5.1
r_b	基区电阻(混合 Π 模型)	7.5
r_j	结半径	9.2
s	表面复合速度	5.2
v_d	漂移速度	1.2
v_l	极限速度	1.2
v_{th}	电子的平均热运动速度	1.2
v_{sat}	饱和(最大)载流子速度	1.2,9.2
x_B	电中性基区宽度(BJT)	5.1
x_E	电中性发射区宽度(BJT)	6.2
x_d	耗尽区宽度	3.2
x_i	"方盒"近似注入深度	9.1
x_j	结深	2.10
$x_{n,p}$	耗尽区边界(n 型区,p 型区)	4.3
x_{ox}	氧化层厚度	2.3
x_{dmax}	最大耗尽层宽度	8.2
α	吸收系数	1.1
α	体电荷因子(MOSFET)	9.1
$\alpha_{F,R}$	正向,反向共基极电流增益	6.4
α_T	基区输运系数(BJT)	6.2
$\alpha_{n,p}$	电子,空穴电离率	4.4
$\beta_{F,R}$	正向,反向发射极电流增益	6.2
γ	发射效率	6.2
γ	体效应系数(MOSFET)	9.1
δ	混合 Π 模型中的亏损因子	7.5
η	效率因子:表面势对栅极电压的微分	9.1
η	混合 Π 模型中的参数	7.5
ϵ	介电常数	1.1
μ_{eff}	MOSFET 沟道中的有效迁移率	9.2

$\mu_{n,p}$	电子,空穴迁移率	1.2
λ	平均自由程	1.2,4.4,10.3
λ_{ph}	光学声子平均自由程	表1.3
$\phi(x)$	x 处的电势	3.3
ϕ_{Bn}	电子注入势垒	4.2
ϕ_{Bp}	空穴注入势垒	4.2
ϕ_{bs}	体到表面的能带弯曲量	8.3
ϕ_{fi}	本征 Fermi 势	1.1
$\phi_{fn,fp}$	电子,空穴的准 Fermi 势	1.1
ϕ_i	内建电势	3.2
$\phi_{p,n}$	p 型区,n 型区的电势	4.2
ϕ_s	表面势	9.1
ρ	电阻率	1.0
ρ	空间电荷浓度	3.4
σ	电导率	1.2
$\sigma_{n,p}$	俘获截面(电子,空穴)	5.1
τ	失效时间	10.4
τ	初始氧化厚度参数	2.3
τ_0	少数载流子寿命	8.4
τ_A	Auger 寿命	5.2
τ_B	基区渡越时间	7.3
$\tau_{BF,BR}$	正向,反向基区渡越时间	7.4
$\tau_{F,R}$	正向,反向电荷控制时间常数	7.3
τ_c	两次碰撞间的平均时间	1.2
τ_{inv}	形成反型层的时间	8.4
$\tau_{n,p}$	过剩电子,空穴寿命	5.2
τ_{nA}	Auger 复合寿命	5.2
τ_r	介电弛豫时间	3.6
$\Gamma_{n,p}$	电子,空穴的 Auger 系数	5.2
ΔE_c	导带边不连续量	4.2
ΔE_g	带隙不连续量	4.2
ΔL	夹断区宽度	9.1
ΔR_p	离子注入的标准偏差	2.5,9.1
$\Phi_{M,S}$	金属,半导体的功函数	8.1
Θ	隧穿几率	4.4
χ	电子亲和能(势)	3.2



索　引

A

Abrupt pn-junction nomograph,突变 pn 结列线图

Absorption coefficient for photons,光吸收系数

Acceptor,受主

Aluminum metallization,铝金属化

Analog-digital converter,模/数转换器

Auger recombination,Auger 复合

Avalanche breakdown,雪崩击穿

Avalanche injection into silicon dioxide,雪崩注入二氧化硅

B

Band diagram,energy-band diagram,能带图

Berglund's method,Berglund 方法

Bipolar junction transistor (BJT),双极晶体管

active bias,放大偏置

amplifying transistor,放大晶体管

bandgap grading,带隙缓变

base resistance,基区电阻

base transit time,基区渡越时间

base transport factor,基区输运系数

buried layer,埋层

charge-control model,电荷控制模型

CMOS latch-up,CMOS 闩锁效应

computer simulation model,计算机建模

current gain,电流增益

defect factor,缺损因子

diffusion capacitance,扩散电容

diode connections,二极管联结方式

Early effect,Early 效应

Early voltage,Early 电压

Ebers-Moll model,EM 模型

emitter efficiency,发射效率

emitter follower circuit,射极跟随器电路
emitter injection,发射区注入
emitter push effect,发射区推进效应
frequency response,频率响应
Gummel Number,Gummel 数
Gummel-Poon model,GP 模型
Heterojunction (HBT),异质结双极晶体管
high-level injection,大注入
hybrid-pi equivalent circuit,混合 Π 等效电路
junction capacitance,结电容
Kirk effect,Kirk 效应
Lateral pnp transistor,横向 pnp 晶体管
linking current,连接电流
LOCOS process,局部氧化工艺
pnp transistors,pnp 晶体管
polysilicon emitter,多晶硅发射极
prototype,原型晶体管
punchthrough,穿通
regions of operation,工作区
saturation transient,瞬态饱和
saturation,饱和
Schottky-clamping,Schottky 箝位
small-signal diffusion capacitance,小信号扩散电容
small-signal model,小信号模型
substrate pnp transistors,衬底 pnp 晶体管
switching,开关
transconductance,跨导
transistor action,晶体管工作原理
walled emitter,墙围发射区
washed emitter,泡发射区
Webster effect,Webster 效应
Body-bias effect (MOSFET),体偏置效应
Bond model of crystals,晶体成键模型
Bonding,flip-chip,键合,倒装焊
Boundary values for minority carriers,少数载流子的边界条件
Breakdown,击穿
avalanche,雪崩
pn junction diode,pn 结二极管
voltage,one-sided step junction,电压,单边突变结

Zener breakdown, Zener 击穿
BSIM MOSFET model, BSIM MOSFET 模型
Built-in voltage at pn junction, pn 结内建电压
Buried layer, 埋层

C

Capacitance 电容
Gauss' law, Gauss 定理
diffusion BJT, 扩散 BJT
diffusion, pn junction, 扩散, pn 结
MOS system, 见 MOS 系统
Capture cross section, 俘获截面
Channel, 沟道
Charge-control analysis BJT, BJT 的电荷控制分析
Charge-coupled devices, 电荷耦合器件
Chemical mechanical polishing (CMP), 化学机械抛光
Chemical vapor deposition, 化学气相淀积
amorphous silicon, 非晶硅
low-pressure (LPCVD), 低压 CVD
insulating films, 绝缘薄膜
plasma enhanced (PECVD), 等离子增强
polycrystalline silicon, 多晶硅
reaction kinetics, 反应动力学
Clean room, 超净间
CMOS (Complementary MOS), 互补 MOS
design, 设计
guard rings, 保护环
latch-up, 闩锁效应
punchthrough, 穿通
twin tubs, 双阱
well depth, 阱深
Compound semiconductors processing, 化合物半导体工艺
Continuity equation, 连续性方程
Copper metallization, 铜金属化
Crystal growth, 晶体生长
Czochralshi silicon crystal growth, 直拉法生长单晶硅

D

Damascene process, Damascene 工艺

Debye length,Debye 长度
Defect factor,缺损因子
Density of states function,态密度分布函数
Depletion approximation,耗尽近似
Depletion-layer width maximum,最大耗尽层宽度
Deposition CVD,化学气相淀积
Deposition gaseous,气相淀积
Detailed balance,细致平衡
Diamond lattice,金刚石结构
Dielectric relaxation time,介电弛豫时间
Diffused resistor,扩散电阻器
Diffusion (of impurities into silicon),(杂质在硅中的)扩散
  concentration dependence,浓度关系
  diffusivity,扩散系数
  dopant redistribution,杂质再分布
  gaseous deposition,气相淀积
  oxidation-enhanced,氧化增强
  polycrystalline-silicon,oxide doping sources,多晶硅、氧化物掺杂源
  profiles for BJT,双极晶体管的杂质分布
  segregation,分凝
  simulation methods,模拟方法
  solid solubility,固溶度
Diffusion capacitance(BJT),扩散电容(BJT)
Diffusion coefficient,扩散系数
Diffusion current,扩散电流
Diffusion length,扩散长度
Diode,二极管,见 pn 结
  BJT connections,双极晶体管的连接方式
  breakdown,击穿
  clamps,箝位
  gated diode,栅控二极管
  generation current in reverse bias,反向偏置产生电流
  heterojunction,异质结
  ideal analysis,理想二极管分析
  integrated circuit,集成电路二极管
  light-emitting,efficiency,光发射,效率
  long-and short-base,长二极管和短二极管
  long-base,长二极管
  photodiode,发光二极管

photo emitting gallium nitride,光发射氮化镓
Schottky barrier,Schottky 势垒二极管
short-base,短二极管
transient behavior,瞬态特性
turn-on voltage,开启电压
Donor,施主
Dopant,杂质
addition and diffusion,掺杂和扩散
compensation,补偿
density versus resistivity,浓度与电阻率关系
donors and acceptors,施主和受主
graded impurity distribution,缓变杂质分布
heavy doping,重掺杂
ion implantation,离子注入
redistribution,再分布
segregation,分凝
Drain induced barrier lowing (DIBL),漏致势垒降低
Drift velocity,漂移速度
Dynamic random-access memory (DRAM),动态随机存储器

E

Early effect,Early 效应
Ebers-Moll model,E-M 模型
Effective mass,有效质量
Einstein relation,Einstein 关系
Electromigration,电迁移
Electron affinity, 电子亲和能
Emitter push (BJT),发射极推进
Energy levels of elemental impurities in silicon,硅中杂质能级
Energy-band diagram basic theory,能带图基本概念
graded impurity distribution,缓变掺杂分布
heavy doping effect,重掺杂效应
heterojunction,异质结
metal-semiconductor system,金属-半导体系统
metal-oxide-semiconductor,金属-氧化物-半导体
thermal-equilibrium statistics,热平衡统计分布
Epitaxy,外延
Etching,刻蚀
dry,干法刻蚀

reactive ion,反应离子
selective,选择性
wet,湿法腐蚀

F

Fermi level,Fermi 能级
quasi-Fermi level,准 Fermi 能级
Fermi-Dirac distribution function,Fermi-Dirac 分布函数
Fick's second law,Fick 第二定律
Field oxide,场氧
Flip-chip bonding,倒装焊
Forbidden-energy gap,禁带宽度
Free electron energy,自由电子能量

G

Gallium arsenide processing,砷化镓工艺
Gallium arsenide properties,砷化镓特性
Gate oxide thickness measurement,栅氧厚度测量
Gated-diode structure,栅控二极管结构
Gauss's law,Gauss 定理
Generation-recombination,产生与复合
Auger recombination,Auger 复合
capture and emission,俘获和发射
capture cross section,俘获截面
electron emission from trap,电子从陷阱发射
excess carrier lifetime,过剩载流子寿命
hole capture by trap,陷阱俘获空穴
localized states,局域态
recombination center,复合中心
Shockley-Hall-Read theory,SHR 理论
surface effects on pn junctions,pn 结表面效应
surface recombination,表面复合
surface-recombination velocity,表面复合速率
Germanium,锗
Graded impurity distributions,缓变杂质分布
complementary error function,余误差分布
Gaussian,Gauss 分布
Gummel number,Gummel 数
Gummel-Poon model,Gummel-Poon 模型

H

Hall effect,霍尔效应

Heterojunction,异质结

bipolar transistors,双极晶体管

currents,电流

charge storage,电荷存储

energy-band notch,能带凹槽

energy-band spike,能带尖峰

thermionic emission,热发射

tunneling,隧穿

Hole,空穴

I

Integrated circuit (IC),集成电路

buried layer,埋层

feature size,特征尺寸

resistor,电阻器

sheet resistance,方块电阻

interconnection and packaging compound semiconductors,化合物半导体的互连和封装

interconnection and packaging,互连和封装

interconnection delay,互连延迟

LOCOS process,局部氧化工艺

process model,工艺模型

Interface charge,界面电荷

International technology roadmap for semiconductors (ITRS),国际半导体技术发展蓝图

Intrinsic silicon,本征硅

Ion implantation,离子注入

approximate "box" distribution,方盒近似分布

CMOS,互补 MOS

dose,剂量

implant straggle,注入标准偏差

ion channeling,离子注入沟道效应

ion range,离子注入射程

MOSFET threshold voltage adjustment,阈值电压调整

simulation,模拟

Ionization coefficient,电离率

Isolation,隔离

junction,结

oxide,氧化物
trench,沟槽

J
JFET,结型场效应晶体管

K
Kirk effect (BJT),Kirk 效应(BJT)

L
Lawrence and Warner analysis (pn junctions),Lawrence and Warner 分析法(pn 结)
Lifetime,寿命
Lithography,光刻
Local oxidation process (LOCOS),局部氧化工艺
Localized states,局域态
capture and emission,俘获和发射
Long-base diode,长二极管
Low-level injection,小注入

M
Mass-action law,质量作用定律
MESFET,结型场效应晶体管
Metallization,金属化
Metal-oxide-silicon (MOS) system,金属-氧化物-半导体系统
band diagram,能带图
capacitance,电容
charge-coupled devices,电荷耦合器件
C-V behavior,C-V 特型
C-V practical considerations,C-V 测量的实际问题
electronics,电学特性
energy-band diagram,能带图
flat-band voltage,平带电压
gate-electrode materials,栅电极材料
ideal,理想状态
memories,存储器
models,模型
non-ideal,非理想状态
oxide and interface charge,氧化层和界面电荷
parameter extraction,参数提取

polysilicon gate,多晶硅栅
small-signal capacitance,小信号电容
surface accumulation,表面积累
surface depletion,表面耗尽
surface inversion,表面反型
surface-charge conditions,表面电荷状态
trench capacitor,沟槽电容器
Metal-Semiconductor Contacts 金属-半导体接触
capacitance,电容
energy-band diagram,能带图
Mott barrier,Mott 势垒
Ohmic contacts,欧姆接触
Schottky barrier,Schottky 势垒
Thermal-equilibrium surface effects,热平衡表面效应
Minority-carrier storage,少数载流子存储
junction and free-carrier storage 结和自由载流子存储
transient behavior,瞬态特性
Mobility,迁移率
scattering,散射
holes and electrons,电子和空穴
temperature dependence,温度关系
Moore's law,Moore 定律
MOS system,MOS 系统
MOSFET,MOS 晶体管
basic behavior,基本原理
body-bias effect,衬偏效应
BSIM3v3 model,BSIM3v3 模型
bulk charge effect, 体电荷效应
channel stops,沟道阻止
channel-length modulation,沟道长度调制
complementary (CMOS),互补型(CMOS)
degradation mechanisms,退化机制
depletion mode,耗尽型
double diffused drain,双扩散漏区
drain-saturation current,漏区饱和电流
drain-saturation voltage,漏区饱和电压
effective mobility measurement,有效迁移率测量
enhancement model,增强型
equations (table),公式(表)

gate coupling,栅耦合
gate current,栅极电流
gate current from drain avalanche hot carriers,漏区雪崩热电子栅极电流
gate current,p-channel devices,栅极电流,p 沟器件
gate current lucky-electron model,栅极电流幸运电子模型
gate-oxide thickness,栅氧化层厚度
gradual-channel approximation,缓变沟道近似
hot carrier injection,热载流子注入
ideality factor,理想因子
lifetime testing,寿命测试
lightly doped drain,轻掺杂漏区
long channel,长沟道
maximum channel field,最大沟道电场
nonvolatile memory structures,不挥发存储器结构
parasitic bipolar transistor,寄生双极晶体管
physical effects and models. 物理效应和模型
pinch-off,夹断
pseudo two-dimensional model for velocity saturated region,速度饱和区准二维模型
punchthrough,穿通
reverse body-bias effect,反衬偏效应
scaling,按比例缩小
small-signal equivalent circuit,小信号等效电路
substrate current,衬底电流
subthreshold conduction,亚阈值区导电
symbols,电学符号
threshold voltage adjustment,阈值电压调整
threshold voltage measurement, 阈值电压测量
transconductance,跨导
velocity-saturation,速度饱和
water analogy,水库模型
MOSFET short-channel effects,MOSFET 的短沟效应
drain current saturation,漏极电流饱和
drain induced barrier lowering (DIBL),漏致势垒降低
frequency response,频率响应
inverse narrow-width effect,反窄沟效应
mobility degradation,迁移率下降
output resistance,输出电阻
scaling,按比例缩小
source-drain charge sharing,源漏电荷共享

speed,速度
subsurface punchthrough,次表面穿通
transconductance,跨导
velocity saturation,速度饱和

N

Nonrectifying (Ohmic) contact,非整流(欧姆)接触
Numerical techniques,数值方法
Gummel's method,Gummel 法
Newton's method,牛顿法

O

Oxidation,氧化
chlorine oxidation,掺氯氧化
concentration-enhanced oxidation,杂质增强氧化
high-permittivity oxides,高 κ 氧化层
linear, parabolic rate coefficients,线性,抛物线速率常数
local oxidation (LOCOS),局部氧化
nitrided oxides,氮氧化硅
simulation,模拟
Oxide charge,氧化层电荷
Oxide isolation,氧化物隔离
Oxide thickness,氧化层厚度

P

Photoconduction,光电导
pin diode,pin 结构二极管
Pinch resistor,夹断电阻
Pinch-off (JFET),夹断(JFET)
PISCES simulation,PISCES 模拟
Pixel,像素
Planar process,平面工艺
pn junction,pn 结
boundary values,边界条件
breakdown,击穿
charge storage,电荷存储
continuity equation,连续性方程
current flow,电流
current-voltage characteristics,电流电压关系

diffusion capacitance,扩散电容
diode transients,二极管瞬态效应
exponential doping,指数掺杂分布
field-effect transistor (JFET),场效应晶体管
gated diode,栅控二极管
generation current in reverse bias,反偏产生电流
heterojunction,异质结
ideal-diode analysis,理想二极管分析
Lawrence and Warner analysis,Lawrence and Warner analysis 分析
linearly graded junction,线性缓边结
long-base diode,长二极管
low-level injection,小注入
lumped-element model,集总元件模型
modeling and simulation,建模和模拟
minority-carrier storage,少数载流子存储
nomograph, 列线图
photodiodes,发光二极管
potential barrier,势垒
quasi-neutrality approximation,准中性近似
reverse bias capacitance,反偏电容
reverse bias maximum field,反向最大电场
reverse-biased step junction,反偏突变结
saturation current, 饱和电流
Shockley-Hall-Read (generation-recombination) theory,SHR(产生复合)理论
short-base diode,短二极管
small-signal equivalent circuit, 小信号等效电路
space-charge-region currents,空间电荷区电流
step (abrupt) junction,突变结
surface effects,表面效应
transient behavior, 瞬态特性
pnp transistor (BJT),pnp 晶体管
Polycrystalline silicon,多晶硅
Polysilicon gate,多晶硅栅
nonequilibrium,非平衡
surface charge,表面电荷
surface potential,表面势
thermal equilibrium,热平衡
Potential,surface,电势,表面
Punchthrough,穿通

Q

Quasi-equilibrium,准平衡

Quasi-Fermi levels,准 Fermi 能级

Quasi-neutral approximation,准中性近似

Quasi-static approximation (charge-control analysis),准静态近似(电荷控制分析)

R

Rate coefficient (oxidation),linear,parabolic,速率常数(氧化),线性,抛物线

Reactive ion etching,反应离子刻蚀

Recombination,复合

　centers,复合中心

　lifetime of excess carriers,过剩载流子寿命

Shockley,Hall,Read theory,SHR 复合理论

　velocity,复合速度

Redistribution (of dopants),(杂质)再分布

Resistivity versus doping density,电阻率与掺杂浓度的关系

S

Silicide process,硅化物工艺

Scaling of IC features,IC 特征尺寸按比例缩小

Schottky barrier lowering,Schottky 势垒的降低

Schottky contacts,Schottky 接触

　barrier heights,势垒高度

　Ohmic contact,欧姆接触

Schottky diode,clamp,Schottky 二极管,箝位

Self-aligned metal silicide (salicide),自对准金属硅化物

Sheet resistance,方块电阻

Shockley,Hall,Read recombination theory,SHR 复合理论

Short-base diode,短二极管

Silicon,硅

　amorphous,非晶硅

　Czochralski,直拉法

　dopant addition and diffusion,掺杂和扩散

　epitaxy,外延

　float zone,区熔

　polycrystalline,多晶硅

　properties,特性

　wafers,硅片(晶圆片)

Simulation and modeling,模拟和建模
BSIM3v3 MOSFET model,BSIM3v3 模型
BJT Gummel-Poon model,双极晶体管 GP 模型
device,器件模拟
lucky-electron model,幸运电子模型
oxidation,氧化模拟
PISCES,器件模拟程序
process,工艺模拟
pseudo two-dimensional model (MOSFET),准二维模型(MOSFET)
Shockley,Hall,Read theory,SHR 复合理论
substrate current,衬底电流
SUPREM,斯坦福大学工艺工程模型
Solubilities of dopants,杂质固溶度
Space-charge-region currents,空间电荷区电流
Step junction,突变结
Stress calculations,应力计算
SUPREM process model,SUPREM 工艺模型
Surface charge,表面电荷
Surface potential,表面电势
Surface recombination,表面复合
Surface states,表面态

T

Thermal equilibrium,热平衡
Thermal velocity,热运动速度
Threshold voltage,阈值电压
Transconductance,跨导
Transistor action,晶体管作用
Trapping,陷阱
Trench capacitor,沟槽电容
Trench isolation,沟槽隔离
Tungsten metallization,钨金属化
Tunneling,隧穿

V

Valence band,价带
Velocity of carriers,载流子速度

W

Work function,功函数

Z

Zener breakdown,Zener 击穿

Zincblende lattice,闪锌矿结构

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝集成电路器件电子学
作者＝
页数＝１０００
ＳＳ号＝０
出版日期＝